

Essays in Applied Microeconomics

Dissertation
for the Faculty of Economics, Business
Administration and Information Technology of the
University of Zurich

to achieve the title of
Doctor of Philosophy in Economics

presented by
Beatrice Brunner
from Wald ZH

approved in July 2013 at the request of

Prof. Dr. Josef Zweimüller
Prof. Dr. Rainer Winkelmann

The Faculty of Economics, Business Administration and Information Technology of the University of Zurich hereby authorizes the printing of this Doctoral Thesis, without thereby giving any opinion on the views contained therein.

Zurich, July 17, 2013

Chairman of the Doctoral Committee: Prof. Dr. Dieter Pfaff

Contents

| | |
|--|-----|
| PREFACE | VII |
| CHAPTER 1: INTRODUCTION | 1 |
| CHAPTER 2: THE IMPACT OF LABOR MARKET ENTRY CONDITIONS ON INITIAL JOB ASSIGNMENT AND WAGES | 7 |
| 2.1 Introduction | 7 |
| 2.2 Data and Sample | 11 |
| 2.2.1 Data Source | 11 |
| 2.2.2 Key Measures | 11 |
| 2.2.3 Sample Selection | 12 |
| 2.2.4 Sample Description | 13 |
| 2.3 The Persistence of Initial Labor Market Shocks | 15 |
| 2.3.1 Graphical Evidence on the Evolution of Wages and Initial Con- ditions | 15 |
| 2.3.2 Econometric Framework | 17 |
| 2.3.3 Main Results: Initial Labor Market Conditions and Wages . . . | 19 |
| 2.3.4 Robustness | 21 |
| 2.3.5 Endogenous Labor Market Entry | 23 |
| 2.3.6 Quantitative Implications | 28 |
| 2.4 Initial Job Assignment | 29 |

| | | |
|-------|---|----|
| 2.4.1 | Quality of Workers' First Employer | 29 |
| 2.4.2 | The Importance of the Initial Employer for the Evolution of Wages | 33 |
| 2.5 | Job Mobility and Quality of Subsequent Employers | 35 |
| 2.6 | Conclusions | 38 |
| 2.A | Appendix: Sample Construction | 39 |

CHAPTER 3: ANNOUNCEMENT EFFECTS OF HEALTH POLICY REFORMS: EVIDENCE FROM THE ABOLITION OF AUSTRIA'S BABY BONUS

43

| | | |
|-------|---|----|
| 3.1 | Introduction | 43 |
| 3.2 | The Austrian Baby Bonus | 46 |
| 3.2.1 | Institutional Background | 46 |
| 3.2.2 | The Abolition of the Baby Bonus | 47 |
| 3.2.3 | The Window of Opportunity | 47 |
| 3.3 | Data and Descriptives | 48 |
| 3.3.1 | Data Source | 48 |
| 3.3.2 | Sample Period(s) | 49 |
| 3.3.3 | The Monthly Birth Count, 1990–2006 | 49 |
| 3.4 | Fertility Response | 51 |
| 3.4.1 | Estimating the Fertility Response | 51 |
| 3.4.2 | Robustness | 54 |
| 3.4.3 | Conception (Re-)Scheduling vs. Timing of Births | 56 |
| 3.5 | Taking Risks for the Bonus? | 58 |
| 3.5.1 | Birth Complications | 58 |
| 3.5.2 | Newborns' Health | 60 |
| 3.6 | Conclusions | 62 |
| 3.A | Appendix: Decomposition | 63 |

CHAPTER 4: CULTURAL DISTANCE AND ATTITUDES TOWARDS IMMIGRATION: EVIDENCE FROM SWISS VOTING RESULTS

65

| | | |
|-------|---|----|
| 4.1 | Introduction | 65 |
| 4.2 | Cultural Values and Beliefs | 69 |
| 4.3 | Data and Descriptives | 71 |
| 4.3.1 | National Votes on Immigration | 71 |
| 4.3.2 | Swiss Population Census | 72 |

| | | |
|-------|--|----|
| 4.3.3 | Descriptives | 75 |
| 4.4 | Empirical Model and Identification | 78 |
| 4.4.1 | Econometric Framework | 78 |
| 4.4.2 | Validity of the Identifying Assumption | 80 |
| 4.5 | Results | 82 |
| 4.5.1 | Main Results | 82 |
| 4.5.2 | Robustness | 84 |
| 4.5.3 | The Role of the Voter Turnout | 90 |
| 4.5.4 | The Rise of Switzerland's Right-Wing Party | 92 |
| 4.6 | Conclusions | 94 |
| 4.A | Appendix | 96 |

| | |
|--------------|-----|
| BIBLIOGRAPHY | 101 |
|--------------|-----|

| | |
|------------------|-----|
| CURRICULUM VITAE | 109 |
|------------------|-----|

Preface

“I may not have gone where I intended to go, but I think I have ended up where I needed to be.”

Douglas Adams (1952–2001), English writer, humorist and dramatist

Writing my doctoral thesis has been a challenging but also a very enjoyable experience. I would like to thank everyone involved for the intellectual, moral, and financial support. In retrospect, I was lucky to have made the right choices and have met inspiring people who motivated me to change the course of my career as an economist. Rafael Lalive’s course on causal analysis aroused my interest in modern econometric methods and convinced me to write my term paper at the Chair of Macroeconomics and Labor Markets. It is at this chair where I got introduced to the fascinating world of labor economics, and, most importantly, where my passion for programming and my particular interest in econometric methods could thrive in combination.

First and foremost, I would like to thank my supervisor Josef Zweimüller for giving me the unique opportunity to write this dissertation at his chair and for his continued support and encouragement. Josef created (and still creates) a highly inspiring and pleasant working atmosphere while allowing greatest possible scientific freedom. He taught me how to frame a promising research project and always helped me to exploit every project’s full potential. I am similarly indebted to Andreas Kuhn for his great guidance and unfailing support. His unlimited zeal and his infectious enthusiasm for research have been driving forces throughout this project. He taught me how to conduct sound and convincing empirical research, and he thereby strongly shaped my own way of thinking about empirical research questions. Many thanks go to Andreas also for providing me with material and links that I could not possibly have

discovered on my own and for patiently answering all sorts of technical questions. It is definitely not a coincidence that he earned himself the nickname “stataprofi”. I cannot emphasize enough how much I have benefitted from Josef’s and Andreas’ knowledge and experience during the pursuit of these studies.

I am also grateful to many teachers, colleagues, and friends. It is impossible to mention each of them by name, but I would like to express my appreciation for their input which has, both directly and indirectly, enhanced the quality of this work. I would like, however, to thank Rainer Winkelmann for being my co-advisor and giving me valuable comments and suggestions when I was presenting my work in his seminar. I am also indebted to Reto Föllmi, Peter Egger and all the participants of the internal seminar for their probing questions and farseeing suggestions which inspired me to examine the subject from different perspectives. I am also grateful for the many helpful comments by the participants of the National Research Network workshops in Linz and Brixton. And I, of course, especially thank Rafael Lalive, whose course on causal analysis made me end up where I needed to be.

Furthermore, I like to express my gratitude to Claudia Bernasconi and Andreas Kohler for their incredibly enjoyable and motivating companionship during the time we shared at the University of Zurich and the Study Center Gerzensee, and for the numerous stimulating discussions we had during the last four years. Further thanks to Philippe Ruh, Jean-Philippe Wüllrich and Christian Hepenstrick for the good time we spent together at the University of Zurich and for regularly organizing social events such as house-warming parties, barbecues and the like. In addition, thanks go to Andreas Steinhauer for valuable discussions and helpful comments, to Simone Büchi for introducing me to the Austrian Social Security Database, and to all people that contributed to making the Chair of Macroeconomics and Labor Markets a great place to work. This includes Sandro Favre, Simone Gaillard, Patrick DeCaes, Aderonke Osikomino, Christian Kiedeisch, Andreas Kettemann, Tobias Renkin, Katrin Koller and all the student assistants.

A very special thank goes to Andreas for always believing in me and staying with me through all the ups and downs. Thank you for enriching my life in so many ways. Our challenging outdoor activities perfectly balance day-to-day research. Special thanks also go to Magi, Heinz, Daniela and Jürg Vollenweider for letting me become part of their family and always giving me a sense of safety and stability. And last but not least, I like to thank my parents, Christine and Hans-Rudolf, for always believing in me. This book is dedicated to all of them.

Beatrice Brunner, Zürich, December 2012

CHAPTER 1

Introduction

This dissertation consists of three self contained essays in applied microeconomics. While disparate in their subject, they share the common theme of rigorous econometric identification with the aim of exploring interesting scientific questions and to test policy relevant hypotheses in various research areas. Chapter 2 belongs to the field of labor economics and studies the short and long run impact of labor market entry conditions on job assignment and wages. Chapter 3 adds to the field of health economics. It is concerned with the identification of short-term fertility and health effects following the early announcement of a health policy reform. Chapter 4 contributes to the field of political economy and offers new insights about how attitudes towards immigration are causally affected by the size and cultural composition of the local immigrant population.

The crucial ingredient required in order to provide convincing answers to these research questions is a sound empirical strategy. Therefore, based on the nature of data, each essay adopts a specific econometric technique that is considered to be best suited to explore the effect of interest. Chapter 2 uses both ordinary least squares (OLS) and instrumental variable (IV) methods, and provides evidence that implies that OLS estimates represent a lower bound of the true effect while IV estimates represent an upper bound. The estimation in chapter 3 relies on predetermined longitudinal regularities in the outcome variable. In particular, we make an out-of-sample prediction of the outcome in the absence of the announcement of the policy change and compare it with the actually observed outcome. The difference represents our estimate of the impact of the health policy reform. Chapter 4 also uses instrumental variable tech-

niques. The remainder of this introductory chapter provides a short summary of each of the three essays.

THE IMPACT OF LABOR MARKET ENTRY CONDITIONS ON INITIAL JOB ASSIGNMENT AND WAGES

The recent economic crisis has renewed academic interest in the potential impact of business cycle fluctuations on labor markets. However, while there exists extensive evidence on the short-run association between local labor market conditions and real wages, only few evidence exists on the longer-run effects of business cycle fluctuations. Clearly, even small initial wage shortfalls may, in the longer run, accrue to substantial overall losses in lifetime earnings if initial wage losses resulting from poor entry conditions persist.

In the first chapter, we estimate the short and long-run impact of initial labor market conditions on wages for young males entering the Austrian labor market between 1978 and 2000. Additionally, we analyze changes in the quality of workers' first employer over the business cycle and the importance of first employers for individuals' entry and subsequent wages. We do so using Austrian social security records that contain detailed individual earnings and employment histories for the universe of private-sector employees from 1972 to 2005. We restrict our sample to low and medium-skilled workers, more precisely, we focus on individuals with no more than secondary education (i.e. vocational training). Our measure for external initial labor market conditions is the annual male unemployment rate at the state-level.

As relates to our estimation strategy, we apply both OLS and instrumental variable methods. While focussing on OLS, we also take into account that the composition of labor market entrants may be endogenous with respect to variations in the local business cycle, and instrument the unemployment rate at labor market entry with the unemployment rate at age 16, the age at which mandatory schooling ends. Our results suggest that IV estimates provide an upper bound of the wage effects, whereas direct evidence on positive sample selection in economic downturns implies that OLS estimates represent a lower bound of the wage effects of local entry conditions.

Consistent with previous evidence, we find a substantial wage penalty from poor entry conditions on starting wages. Our preferred estimates imply that an increase in the initial unemployment rate by one percentage point is associated with 0.9% lower starting wages. Moreover, this initial wage loss from entering the labor market during high unemployment turns out to be highly persistent, resulting in sizeable negative effects on lifetime earnings. An increase in the initial unemployment rate by one percentage point is estimated to result in a lifetime loss in wages of about 1.3%. We

further show that there is considerable cyclical variation in the quality of a worker's first regular employment. Even though workers are positively selected in times of high unemployment, they end up working in firms that are younger and smaller, and that generally pay lower wages. Moreover, we show that the characteristics of workers' first employers explain as much as three-quarters of the observed long-run effects resulting from poor entry conditions. Finally, we find some very distinct patterns for blue and white-collar workers. While the negative wage effects fade away for white-collar workers after about five to ten years, the wage effects for blue-collar workers are much more persistent because some of them appear to be permanently locked in into low-paying jobs.

ANNOUNCEMENT EFFECTS OF HEALTH POLICY REFORMS: EVIDENCE FROM THE ABOLITION OF AUSTRIA'S BABY BONUS

This chapter is motivated by the fact that anticipatory behavior of individuals has potentially important implications for the implementation, as well as for the evaluation, of health policy reforms. More specifically, we may be especially concerned about potential negative health effects if individuals respond to anticipated or announced changes in financial incentives but are unable to foresee all possible health effects resulting from behavioral changes.

In chapter 3, we study the fertility effects, as well as the potential health consequences for both mother and newborn child, following the announcement of the abolition of the Austrian baby bonus as of January 1, 1997. The Austrian baby bonus amounted to a maximum of 1090 € per child and was paid conditional on medical examinations of both mother and child. The unique feature of this policy change is that the elimination of the benefit was announced about ten months prior to enactment. Thus, although the response window was only limited to three weeks, the early announcement created the potential for an announcement effect since prospective parents had both an incentive and the opportunity to react and move their baby plans forward.

Because there is no control group available, our estimation strategy relies on regularities in the data before the announcement. More precisely, we make an out-of-sample prediction of the outcome in the absence of the early announcement, the difference between the observed and the predicted number (i.e. the residual) then represents our estimate of the announcement impact.

We find robust evidence that, within the month before the abolition (December 1996), about 8% more babies were born than in the absence of the announcement of the policy change. Considering the fact that the window of opportunity was quite a short

period of three weeks only, this is a sizeable effect. Moreover, we find (re-)scheduling of conceptions rather than birth timing (through medical intervention methods like cesarean sections) to be the source of the fertility effect. The analysis of birth procedures further reveals a significant increase in the fraction of mothers experiencing some kind of birth complications by about one percentage point. We calculate that only a small fraction of this increase can be attributed to changes in observable maternal characteristics. We argue that some part of the unexplained increase in birth complications is potentially caused by an underlying increase in behavioral and/or psychological risks triggered by (re-)scheduled pregnancies. However, the increase in birth complications notwithstanding, we do not find any adverse immediate impact on newborns' health.

On a more general level, the results of chapter 3 illustrate that announcement effects may be an important issue in health policy reforms. The abolition of the Austrian baby bonus clearly shows that even relatively small changes in financial incentives may trigger substantial behavioral responses. Policy makers should thus be aware that not only a policy (reform) itself, but also the public announcement of its abolition (or introduction) may have an impact on individual behavior.

CULTURAL DISTANCE AND ATTITUDES TOWARDS IMMIGRATION: EVIDENCE FROM SWISS VOTING RESULTS

Chapter 4 is motivated by the fact that global international migration is increasing not only in scale but also in the cultural diversity of groups involved in that movement. As a consequence, the host countries have to cope with increased levels of social and cultural diversity and thus face new challenges that may not only lead to additional costs but also cause social tensions or even stir up anti-immigration sentiments.

In the final chapter of my dissertation, I shed light on the impact of immigration on anti-immigration attitudes from a novel prospective. First, instead of using survey data, I take advantage of Switzerland's system of direct democracy, where citizens vote regularly on immigration issues. I transform these voting results into a share of anti-immigration votes and use them as a direct measure of attitudes towards immigration. Second, I focus on understanding how the cultural distance between the immigrant population and the Swiss affects this relationship. Drawing on findings from the World Value Surveys, I categorize Swiss immigrants into different groups according to cultural values and beliefs of their source country. In particular I differentiate between nationalities with a value system similar to the Swiss (culturally similar immigrants) and nationalities without one (culturally different immigrants).

In summary, I investigate whether and to what extent natives' attitudes towards immigration are causally affected by the culture-specific immigrant shares of the locality in which they live.

In order to explore this question, I combine census data with outcomes of national votes about immigration in Switzerland between 1970 and 2010. To take endogeneity of the place of residence of both natives and immigrants into account, I instrument the immigrant share in the municipality with the immigrant share in the local labor market. Identification therefore relies on the assumption that an individual's mobility is geographically limited by exogenous factors like family and friends. Indeed, the analysis of residential mobility effects in response to immigration provides evidence that strongly supports the validity of the identifying assumption.

The main finding of this chapter is that it is not the overall immigrant share per se, but merely the presence of culturally different immigrants that induces Swiss citizens to vote for a more restrictive immigration policy. In fact, the local share of culturally different immigrants turns out to be a significant and sizeable determinant of the proportion of anti-immigration votes. Baseline estimates suggest that a one percentage point increase in the share of culturally different immigrants results in a 1.12 percentage point increase in the share of anti-immigration votes. By contrast, the share of culturally similar immigrants does not seem to affect natives voting behavior at all. This result holds for each single vote, is not sensitive to a finer cultural classification of immigrants, and does not seem to be driven by selectivity of turnout. Results further suggest a tight link between voting in favor of a more restrictive immigration policy and support for the right-wing Swiss people's party.

These findings may have implications for the continuing debates over immigration policy and help inform policy makers about possible impacts of the changing immigrant composition on outcomes of future votes. Moreover, the finding that the ethnic composition of the locality in which individuals live affects anti-immigration attitudes may also have implications for continuing debates about the nature and location of new refugee accommodation centers.

CHAPTER 2

The Impact of Labor Market Entry Conditions on Initial Job Assignment and Wages

Joint with Andreas Kuhn

A revised version of this chapter is accepted for publication in the *Journal of Population Economics*. Earlier versions have been published in 2009 as IEW Working Paper, No. 457; IZA Discussion Paper, No. 4601; NRN Working Paper, No. 929 and in 2010 as IEW Working Paper, No. 520; IZA Discussion Paper, No. 5360; NRN Working Paper, No. 1015.

2.1 Introduction

The recent economic crisis has renewed academic interest in the potential impact of business cycle fluctuations on labor markets (e.g. Elsby *et al.*, 2010). However, while labor economists have studied the short-run association between local labor market conditions and real wages extensively for quite some time (e.g. Blanchflower and Oswald, 1990), longer-run effects of business cycle fluctuations on individuals' wages have only more recently caught the attention of empirical research.¹ Clearly, even small initial wage shortfalls may, in the longer run, eventually accrue to substantial overall

¹Most studies estimating the short-run association between fluctuations in local unemployment rates and wages find that wages vary negatively with local unemployment. This negative association is a very robust empirical pattern; it has been shown to exist for a wide range of different countries, using very different sources of data and diverse empirical specifications. See Nijkamp and Poot (2005) for a comprehensive survey of this literature.

losses in lifetime earnings if initial wage losses resulting from poor entry conditions persist.²

Indeed, recent empirical evidence suggests that substantial losses in lifetime earnings result from entering the labor market during an economic downturn, as opposed to entering during an expansion. Oreopoulos *et al.* (2012) explore the effects of entering the labor market during a recession on individuals' earnings, using data on Canadian college graduates who entered the labor market between 1982 and 1999. They find a substantial initial wage penalty of about 9% that only fades to zero after the first decade of a worker's career. A similar result is reported in Kahn (2010), who focuses on male college graduates in the United States graduating sometime between 1979 and 1988. She finds that the group graduating in the worst economic situation incurs a wage loss of up to 13% each year, relative to those graduating in the best initial conditions, and that this initial wage loss persists over the first 20 years of a workers' labor market career. Similar results are reported in Oyer (2006), who shows that PhD students in economics are considerably more likely to find a position at one of the top universities in the United States if they graduate in times when the demand for economists is high. In a related study, he finds that those MBA students who complete their training during a recession suffer from negative effects on wages (Oyer, 2008). In both studies, the long-term effects on income appear to stem from the fact that diverse employers hire workers entering the labor market under different conditions, therefore giving them access to distinct jobs. The entry job is thus significant for the future career, and this appears to be of particular importance for highly educated individuals for whom the transition in and out of attractive positions is very low.³ Mansour (2009) also focuses on college graduates in the US and again finds negative and persistent wage effects from entry into the labor force during a recession. Moreover, he shows that failure to account for endogenous sample composition underestimates both the immediate wage effects as well as the persistence of wage effects

²Previous research has shown that the early years in a worker's labor market career are of special importance (Gardecki and Neumark, 1998; Neumark, 2002). In terms of wages, Murphy and Welch (1990) estimate that almost 80% of all (i.e. lifetime) wage increases accrue within the first ten years of labor market experience. Moreover, movements across jobs are considerably more likely at the beginning of a worker's career than later on (Topel and Ward, 1992).

³One important concern regarding the validity of these results is that schooling and first entry into the labor force may be endogenous both because individuals may choose to stay in school or continue further training when faced with high unemployment and low starting wages. Indeed, several studies find that enrollment rates are high when unemployment is high and the opportunity costs of schooling are low (e.g. Clark, 2011). In line with these findings, both Kahn (2010) and Oreopoulos *et al.* (2012) find the duration of schooling to be endogenous. Both tackle the endogeneity problem by instrumenting the unemployment rate at the time of labor market entry with either the prevailing unemployment rate at a lower age or that in the predicted year of graduation. Mansour (2009) presents direct evidence on sample selection over the business cycle based on AFQT scores.

resulting from initial labor market shocks.

Empirical evidence for countries outside the US and Canada yields broadly similar findings. Kwon *et al.* (2010) mainly use data from the Swedish labor market.⁴ They find that workers who enter the labor market during a boom are not only paid higher wages, but that they are also promoted more quickly to higher ranks than those who enter during an economic downturn. Stevens (2007) finds significant negative, albeit small effects of initial conditions on wages in Germany (much smaller than those found in the US and Canada). In contrast to all other studies, however, she finds that wage losses from poor entry conditions do not fade away, but actually increase over time.⁵ The available empirical evidence also underlines the fact that negative wage effects of initial labor market conditions are not confined to highly skilled workers. Genda *et al.* (2010) focus on a separate comparison between men in Japan and the United States with more or less education with respect to the effects of initial conditions. They find negative effects of initial conditions for more highly skilled workers in both countries. However, they only find negative wage effects for workers in Japan with fewer skills. They argue that the specific hiring system and employment protection drive the persistence of the effects for Japanese, while the market for less skilled workers in the United States may indeed be quite close to a competitive market. Consistent with this finding, Kondo (2008) reports that the initial effect of entering the labor market during a recession on wages is less persistent for less skilled workers and for workers with weak labor market attachment in the US.

In this chapter, we present estimates of the long-run effects of business cycle fluctuations on young males' wage profiles in the Austrian labor market and derive an empirical estimate of the associated loss in lifetime earnings due to entering the labor force during a recession, as opposed to entry during average aggregate conditions or during a boom. We do so using Austrian social security records that contain detailed individual earnings and employment histories for the universe of private-sector employees from 1972 until 2005. We complement the available empirical evidence on the long-run wage effects of labor market entry conditions with an analysis for Austria, a labor market characterized by a high level of employment protection and a centralized wage bargaining structure. We focus on low and medium-skilled workers, while most of the studies mentioned above focus on higher or even highest-skilled workers (in

⁴Previous studies for European countries have mainly focused on the long-run effects of initial conditions on employment rather than wages (e.g. Burgess *et al.*, 2003; Raaum and Røed, 2006).

⁵A similar analysis of wage effects for firm entry cohorts in the German manufacturing sector is given in von Wachter and Bender (2008). However, their analysis is not confined to new labor market entrants, but covers workers of all experience levels; their results are therefore not directly comparable to the other studies mentioned.

terms of formal education).⁶ In the second part of the analysis, we focus on changes in the quality of workers' first employer over the business cycle and the importance of first employers for individuals' entry and subsequent wages. Indeed, several previous studies have shown that workers' initial placements may have important effects not only on their entry wages, but also on their subsequent wages (e.g. Oreopoulos *et al.*, 2012; von Wachter and Bender, 2006). More specifically, initial job or task assignment may be important in the longer run if employers assign otherwise identical workers to lower quality jobs or tasks in recession, and if jobs or tasks offer different opportunities for the accrual of human capital (e.g. Gibbons and Waldman, 2006).⁷ Alternatively, workers' initial job or task assignment may have long lasting effects on wages if workers accumulate human capital while on the job that is not fully transferable to other jobs or tasks because it is specific to a worker's task, occupation, or industry. In the third and final part of the empirical analysis we provide complementary evidence on how local entry conditions affect workers' mobility behavior and unemployment experience and how the quality of their subsequent employer changes as they start accumulating labor market experience.

The remainder of this chapter is organized as follows. The next section presents our data source, details the sample selection process, and discusses the construction of our key measures. Section 2.3 presents the econometric approach for estimating long-run wage effects of initial labor market conditions along with our main results. We also present some robustness tests as well as some evidence on the importance of endogenous labor market entry over the business cycle. In section 2.4, we focus on the impact of initial labor market conditions on the quality of workers' first employer, and the importance of first employer on individuals' subsequent wages. Section 2.5 studies how initial conditions affect job mobility and unemployment and how the quality of workers' subsequent employers changes over time. Section 2.6 concludes.

⁶Note, however, that workers in the Austrian labor market typically have some, potentially very specialized, vocational training, as a significant part of the initial vocational training in Austria is provided by dual apprenticeship training schemes, i.e. practical training provided by firms coupled with part-time compulsory attendance at a vocational school. Apprenticeships last from two to four years, depending on occupation. Full-time vocational and technical schools provide an important alternative to apprenticeship training, and also last up to four years. Details are available from the report by the Federal Ministry for Education, the Arts and Culture (2008).

⁷Consistent with this line of argument, Kwon *et al.* (2010) find that workers who enter the labor market during a boom are promoted more quickly and to higher ranks than those who enter during a recession, and Mansour (2009) shows that workers entering in a recession are initially assigned to lower paying jobs.

2.2 Data and Sample

2.2.1 Data Source

We use individual-level social security records from the Austrian Social Security Database (ASSD), a data source described in more detail in Zweimüller *et al.* (2009). The ASSD basically covers the universe of Austrian private sector workers from January 1972 until December 2005 and contains complete and precise information about individuals' annual earnings and daily employment histories. The data are therefore ideally suited for studying the impact of labor market shocks on long-run wage profiles because they allow us to construct individual wage profiles for a large number of labor market entrants over a relatively long period of time.

2.2.2 Key Measures

Our dependent variable for most of the analysis is the real daily wage, i.e. the real wage per actual day of work, adjusted to 2007 prices. Wages are deflated with the consumer price index and include additional/special payments such as a 13th month's salary or holiday pay.⁸ Real daily wages are computed as the average earnings over all employers in a given year. This means that we first sum total annual earnings over all employers for any individual. We then divide overall earnings by the total number of days worked in a given year, also summed over all employers for a given individual and taking overlapping employment spells into account.

The regressor of main interest is the annual male unemployment rate, our measure for external labor market conditions at the time individuals first enter the labor force. We compute annual male unemployment rates from the individual level employment histories contained in the ASSD raw data. This procedure has the advantage that we can calculate unemployment rates back until 1972 (compared to published statistics, which only reach back until 1978) and at different levels of cross-sectional aggregation.⁹

⁸The Austrian Central Social Security Administration collects these data with the purpose of administering and calculating entitlements to old age pension benefits. For this reason, the ASSD includes precise and comprehensive information on annual earnings and daily employment histories. However, contributions to the old age pension system are capped because old age pension benefits are limited to a maximum level. As a consequence, annual earnings are only recorded up to the threshold which guarantees the maximum benefit level ("Höchstbemessungsgrundlage", HBGr). Similarly, there is a base threshold below which no (otherwise mandatory) social security payments accrue ("Geringfügigkeitsgrenze", GfGr). The two censoring points vary over time in real terms: The lower censoring point increased from about 14 € in 1978 to about 26 € in 2005 (per day worked); the upper censoring point increased from about 78 € to 126 € per workday over the same period of time.

⁹We decided to extract yearly male unemployment rates for the age groups 16 to 65 and 16 to 25, both at the state ("Bundesland") level and at the common classification of territorial units for statistics (NUTS) level. At the NUTS level, we use the most disaggregated level available (NUTS-3),

Unless noted otherwise, we use the male unemployment rate for all workers aged between 16 and 65 at the state level as our main regressor.

2.2.3 Sample Selection

Mainly for conceptual reasons, but also due to some data limitations, we do not work with the universe of all labor market entrants but only with a specifically selected sample. First, we restrict our attention to male entrants only. On the one hand, female labor supply behavior over the life cycle is much more difficult to model than male labor supply. On the other hand, we believe that the fact that most men work full-time allows us to largely circumvent the problem that the ASSD does not contain information on working hours. Second, we select those workers who start their first regular employment spell sometime between 1978 and 2000, allowing us to observe at least five additional years of earnings for each worker because the data run until the end of 2005 (see also appendix 2.A). As a final restriction, we focus on workers aged between 16 and 21 at the time they first enter the labor force (i.e. start their first regular employment spell).¹⁰ Essentially, this restriction excludes individuals with higher education (most importantly, individuals with a university degree), but it should include all or most individuals with an apprenticeship training or an education of similar length and scope, such as full-time vocational school.¹¹ Our final sample thus consists of male low and medium-skilled labor market entrants who started their first regular employment between 1978 and 2000. We can observe these workers' full labor market career from the year they first enter into the labor force until the year 2005.

which corresponds to one or more political districts in Austria. There are total of 9 different states and 35 different NUTS-3 regions in Austria. Yearly unemployment rates are within-year averages of monthly unemployment rates.

¹⁰Because the ASSD does not contain a comprehensive measure for schooling, we use age at entry into the labor force as proxy for education in the regressions below. To mitigate potential collinearity with year of birth and year of entry, we use a slightly different variable as proxy in the regressions. Specifically, we use the smaller of age at start of first regular employment and age at start of first registered unemployment spell.

¹¹Several arguments motivate the restriction on schooling. First, the timing of first labor market entry, and thus the duration of schooling, may be endogenous. However, less skilled workers are presumably less likely to manipulate the duration of schooling. Furthermore, unobserved heterogeneity resulting from, say, unobserved differences in inherent ability, is arguably a more urgent problem for workers with higher skills. Moreover, we believe that our proxy for schooling works best for workers with low education levels. Finally, only including less-skilled workers in the sample is an effective way of dealing with right-censored wages (see also appendix 2.A).

2.2.4 Sample Description

Because we can follow all individuals from the year of their first regular employment until the end of the data in the year 2005, the resulting data set would have been too large from a practical point of view. In the following, we therefore work with a 30% random sample of all labor market entrants aged between 16 and 21 when first entering into the labor force. This sample contains 217,587 unique individuals and about 3.35 million individual wage observations (i.e. observations at the level of individual \times year).

Table 2.1 shows descriptive statistics for our final analysis sample. The first panel shows individual-level characteristics. The average labor market entrant in our sample is about 19 years old when starting his first regular employment spell, and he holds his first job for almost three years. The average age at the start of the first job dovetails with the fact that mandatory schooling ends in the year when individuals attain the age of 16 and that apprenticeships usually last for two to four years. The high fraction of blue-collar occupations is consistent with our intention of only (or mainly) including individuals who received some kind of vocational training. Interestingly, a substantial fraction of the sample (about a third) experiences some unemployment before starting the first regular employment spell (these individuals are registered for unemployment benefits on average for somewhat more than one month). Consistent with this, we find that age at first entry (our proxy for schooling) is about half a year lower than age at the start of the first regular job, reflecting the fact that the transition from education to work often involves short periods of non-employment.

The second panel shows various unemployment rates as well as the number of labor market entrants. The different unemployment rates use different aggregation levels (cf. footnote 9) and refer to different age groups, but all of them are limited to the male population only. The unemployment rate in the year of labor market entry averages about 6.6%, irrespective of the chosen aggregation level (states or NUTS-3 units). Youth unemployment rates at the time of entry are somewhat lower than overall unemployment rates and equal about 5% on average. Finally, about 3,750 (2,430) male individuals aged between 16 and 21 (22 and 30) enter the labor market in any given state and year in the period 1978-2000.¹²

The remaining part of the table shows the distribution of our sample of labor market entrants across the nine different states and across years. The distribution across states mainly reflects differences in population size. It may, however, also reflect other differences between states (e.g. differences in the age distribution of

¹²The number of labor market entrants aged between 22 and 30 are used as control variables in some of our model specifications.

Table 2.1: Summary statistics, male labor market entrants 1978–2000

| | Mean | Standard deviation |
|--|-----------|--------------------|
| <i>Individual-level characteristics:</i> | | |
| Age at start of first regular job | 19.195 | (1.007) |
| Age at first entry into the labor force | 18.669 | (1.431) |
| Duration of first regular job (years) | 2.847 | (4.335) |
| Any unemployment before first job | 0.306 | (0.461) |
| Unemployment days before first job | 42.649 | (101.181) |
| Blue-collar | 0.546 | (0.498) |
| White-collar | 0.176 | (0.380) |
| <i>Aggregate-level variables:</i> | | |
| State level unemployment rate | 6.595 | (2.897) |
| NUTS-3 level unemployment rate | 6.645 | (3.144) |
| State level youth unemployment rate | 4.941 | (1.790) |
| Number of entrants aged 16-21 | 3,767.283 | (1,451.150) |
| Number of entrants aged 22-30 ^a | 2,431.488 | (2,070.294) |
| <i>State of first entry into the labor market:</i> | | |
| Vienna | 0.179 | (0.383) |
| Lower Austria | 0.173 | (0.378) |
| Burgenland | 0.025 | (0.156) |
| Upper Austria | 0.194 | (0.395) |
| Styria | 0.147 | (0.354) |
| Carinthia | 0.067 | (0.251) |
| Salzburg | 0.071 | (0.256) |
| Tyrol | 0.092 | (0.289) |
| Vorarlberg | 0.053 | (0.223) |
| <i>Year of first entry into the labor market:</i> | | |
| 1978 | 0.033 | (0.178) |
| 1979 | 0.039 | (0.194) |
| 1980 | 0.046 | (0.208) |
| 1981 | 0.045 | (0.207) |
| 1982 | 0.042 | (0.202) |
| 1983 | 0.044 | (0.206) |
| 1984 | 0.048 | (0.213) |
| 1985 | 0.047 | (0.212) |
| 1986 | 0.046 | (0.210) |
| 1987 | 0.046 | (0.210) |
| 1988 | 0.048 | (0.214) |
| 1989 | 0.050 | (0.217) |
| 1990 | 0.049 | (0.216) |
| 1991 | 0.046 | (0.209) |
| 1992 | 0.042 | (0.201) |
| 1993 | 0.038 | (0.190) |
| 1994 | 0.041 | (0.198) |
| 1995 | 0.040 | (0.196) |
| 1996 | 0.038 | (0.192) |
| 1997 | 0.040 | (0.195) |
| 1998 | 0.041 | (0.197) |
| 1999 | 0.043 | (0.204) |
| 2000 | 0.048 | (0.214) |
| Number of unique individuals | 217,587 | |

Notes: All aggregate-level variables are computed from the individual-level raw data of the ASSD (see footnote 9). See also notes of table A.1. ^a The number of entrants aged between 22 and 30 are used as control variables.

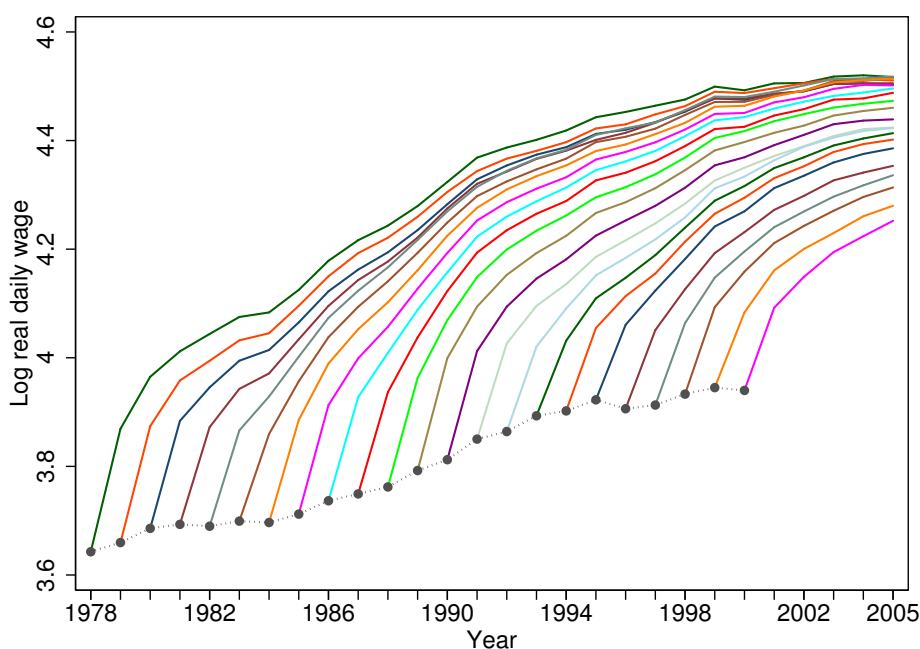
women). The distribution of entrants across years on the other hand looks fairly uniform, implying that the aggregate number of entrants aged between 16 and 21 has not changed much over time.

2.3 The Persistence of Initial Labor Market Shocks

2.3.1 Graphical Evidence on the Evolution of Wages and Initial Conditions

We start with a simple graphical depiction of our two key measures (i.e. cohorts' wage profiles and the initial local unemployment rate). First, figure 2.1 shows wage profiles by entry cohort for all labor market cohorts who first entered the labor force between 1978 and 2000. The black dots therefore represent average starting wages for each entry cohort and the dashed grey line thus shows how starting wages evolve over time. Clearly, real starting wages have increased significantly over the period of analysis, from about 38€ in 1978 to about 50€ in 2000. Also note that there is some cyclical movement in starting wages over time which we expect to be related to economic conditions prevailing in that year.

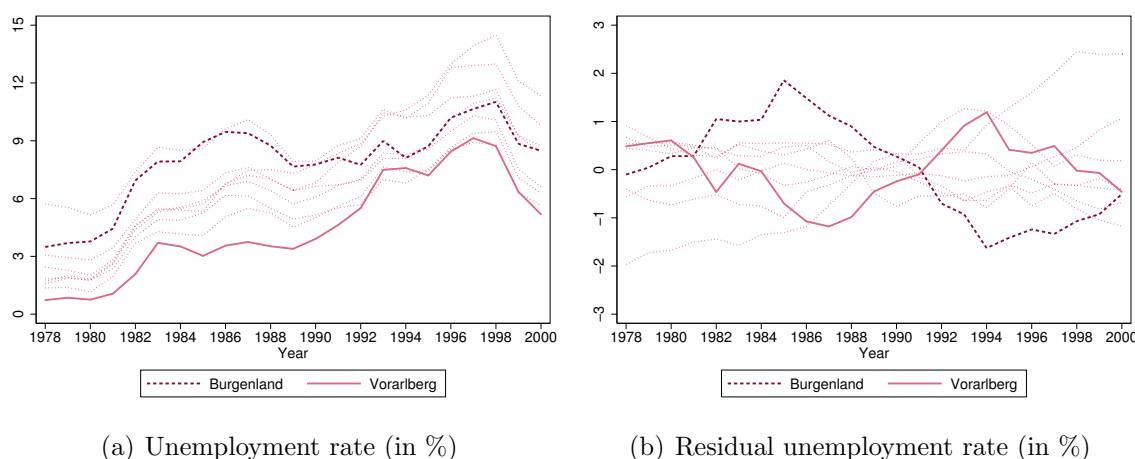
Figure 2.1: Long-run wage profiles, by labor market entry cohort



Notes: The figure shows average log real daily wages by calendar year for each labor market entry cohort from 1978 to 2000.

The solid colored lines, on the other hand, represent long-run wage profiles of cohorts entering the labor market in different years. Cohorts' wages clearly follow an approximate concave path over time, implying that wage growth is highest in earlier working years and then strongly flattens later on. The figure shows, for example, that the 1978 entry cohort starts with a real wage of about 38 € per workday and experiences a raise in real wage up to about 97 € by the year 2005. On average, this cohort's compensation therefore more than doubles in real terms within the first 27 years of labor market experience. Focusing again on the 1978 entry cohort, we see that this cohort's average wage grows by approximately 146% ($= [\exp(.9) - 1] \cdot 100\%$) in the first 27 years of experience. Evidently, most of this wage increase happens at the early stage of the labor market career (i.e. the wage increase in the first ten years equals about 86% ($= [\exp(.62) - 1] \cdot 100\%$)).¹³

Figure 2.2: Fluctuations in state-level unemployment rates, 1978-2000



Notes: The figure shows male unemployment rates for workers aged between 16 and 65, aggregated at the level of the state (“Bundesland”) and computed from the individual-level raw data of the ASSD. Two of the nine states are graphically highlighted. “Vorarlberg” is situated in the western part of Austria (bordering Germany and Switzerland). “Burgenland” is located at the eastern border of Austria (bordering Hungary in the east and Slovenia in the south). Panel (b) shows residuals from a regression of the initial unemployment rate on a full set of year and state of entry dummies.

Panel (a) of figure 2.2 shows the evolution of state level unemployment rates which we use as our indicator for external labor market conditions at the time individuals first enter the labor force. This figure shows that the period from 1978 to 2000 covers

¹³Wage profiles of different entry cohorts have somewhat distinct overall shapes. More specifically, the figure shows that returns to experience generally decrease over time, meaning that younger entry cohorts have considerably lower returns to labor market experience than older cohorts. For example, the 1995 cohort only realizes an average wage increase of about 58% ($= [\exp(.46) - 1] \cdot 100\%$) in the first ten years, thus less smaller than that of the corresponding increase of the 1978 entry cohort.

several periods of both boom and downturn, and that the identifying variation in initial labor market conditions therefore does not only stem from a few neighboring labor market cohorts. The figure also illustrates that states not only differ in the level of unemployment but also with respect to cyclical variations around a longer run trend: although all states see an increase over the whole observation period in general, there are marked cyclical differences across states.¹⁴ Panel (b) of figure 2.2 shows that our observation period spans several ups and downs in the business cycle, and that there is considerable differentiation in the strength of these variations across states. We thus have both sufficient cross-sectional and longitudinal variation in our key regressor that we can use to pin down the effect of local business cycle fluctuations on wages.

2.3.2 Econometric Framework

Because we primarily aim to estimate the long-run impact of economic shocks at the time individuals started their first jobs, we must take care to allow the association between initial conditions and wage to become weaker or stronger as labor market experience increases while also using a generally flexible functional form of wage profiles. Taking these issues into account, our basic and most parsimonious econometric model is the following:

$$\ln(y_{it}) = ur_{j[i]}^0 \alpha_1 + \kappa(\exp_{it}) \alpha_2 + [ur_{j[i]}^0 \cdot \kappa(\exp_{it})] \alpha_3 + \psi_i^{j(t_0)} + \phi_i^{t_0} + \theta_t + \epsilon_{it}, \quad (2.1)$$

where y_{it} denotes the real daily wage of individual i in calendar year t , \exp_{it} potential labor market experience of i in year t , and $ur_{j[i]}^0$ the unemployment rate as prevailing in state j at the time individual i first entered the labor market. Function $\kappa(\cdot)$ denotes that we allow for a flexible functional form with respect to labor market experience.¹⁵ Note that we allow the effect of the initial unemployment rate on current wages to vary as potential labor market experience increases by including the interaction terms between the experience polynomials and the initial unemployment rate. We also include a full set of dummies for state at entry and year of entry, denoted by $\psi_i^{j(t_0)}$ and

¹⁴For example, and as highlighted in the figure, Burgenland (located in southeastern Austria) experienced a huge increase in the unemployment rate from about 3% in the late 1970s to about 8% in the first half of the 1980s, and then to about 9% in the second half of the 1980s. Vorarlberg (situated in western Austria), in contrast, experienced only a modest increase from about 1% in the 1970s to about 3% in the 1980s. In 1992, however, Vorarlberg underwent a sharp decline in the local labor market conditions, when unemployment jumped from about 3% to about 7–8%.

¹⁵Specifically, we include the first three polynomial terms of potential labor market experience. We chose the number of polynomial terms on the basis of a non-parametric, and therefore fully flexible, wage-experience model. The first three polynomial terms appear sufficient for reproducing the wage-experience profile that a corresponding non-parametric specification predicts.

$\phi_i^{t_0}$, respectively, throughout the analysis. Finally, we also control for the aggregate business cycle θ_t which we parameterize using the aggregate yearly unemployment rate, the log aggregate number of workers, as well as log aggregate annual earnings.

As we will discuss in more detail below, our baseline (i.e. our preferred) specification will include some additional regressors:

$$\begin{aligned} \ln(y_{it}) = & ur_{j[i]}^0 \alpha_1 + \kappa(\exp_{it}) \alpha_2 + [ur_{j[i]}^0 \cdot \kappa(\exp_{it})] \alpha_3 + \psi_i^{j(t_0)} + \phi_i^{t_0} + \theta_t \\ & + x_i^0 \beta_1 + [\phi_i^{t_0} \cdot \kappa(\exp_{it})] \beta_2 + \ln(n_{j[i]}^0) \beta_3 + \epsilon_{it}, \end{aligned} \quad (2.2)$$

where the second row contains these additional variables. First, x_i^0 denotes a small set of individual-level controls, i.e. our proxy for schooling and two indicator variables for workers' initial occupation (blue or white-collar). These variables are predetermined in the sense that they relate to an individual's first regular employment spell or to the time before having started to work (i.e. there is no time index for these variables). We also include the full set of interaction terms between year of entry and the polynomial in potential experience as well as the log number of labor market entrants aged 16-21 and aged 22-30 at the state level. The latter two variables are included to control for changes in the relative number of entrants aged 16-21 (relative to all entrants aged in the age bracket 16-30); the former set of controls allows for cohort-specific differences in the wage profile that are unrelated to differences in local initial unemployment rates but due to, say, changes in the production of education across cohorts.

In either case, parameters α_1 to α_3 describe individuals' wage-experience profiles as a function of the initial unemployment rate and are the parameters of main interest. Specifically, α_1 is the elasticity of wages with respect to the initial unemployment rate in the year of first entry (i.e. in the year where labor market experience is equal to 0), while α_3 tells us how the effect of initial conditions changes as labor market experience increases.

One important complication both specifications imply relates to the fact that the local initial unemployment rate does not vary over time for any given individual. For this reason, we cannot use standard panel data estimators such as the fixed-effects or first-differences estimator because these methods not only eliminate all unobserved time-invariant heterogeneity but also all variation in the key regressor. We therefore rely on estimation methods that use the untransformed data. We also have to consider that our key regressor is observed at a higher level of aggregation than the dependent variable, a situation that may lead to grossly misleading statistical inference (Moulton, 1986). All standard errors we report are therefore clustered by cells defined by year at first entry \times state of first entry (there are 9 states and 23 entry years, resulting in

207 distinct cells).¹⁶

2.3.3 Main Results: Initial Labor Market Conditions and Wages

Table 2.2 shows our main results, building up from our most parsimonious specification represented by equation (2.1) to our preferred specification as given by equation (2.2). The top panel shows, for each specification, point estimates of the parameters (i.e. α_1 and α_3) describing the effect of initial conditions on wages at the top and estimated semi-elasticities of wages with respect to the initial local unemployment rate at specific values of potential labor market experience (i.e. potential labor market experience of 0, 5, 10, 15, and 20 years, respectively) in the middle of the table. For example, $\varepsilon_{ur}^y(5)$ denotes the estimated semi-elasticity of the real daily wage with respect to the initial unemployment rate at five years of potential labor market experience. It thus corresponds to the estimated relative change in wages resulting from a one percentage point increase in the initial unemployment rate.

The first column shows point estimates for our most parsimonious specification as spelled out in equation (2.1). As expected, there is a negative effect of the initial local unemployment on wages in the year of entry. Specifically, the semi-elasticity of wages with respect to the initial unemployment rate equals -0.009 in the year of entry. A one percentage point increase in the local initial unemployment rate is thus predicted to lower starting wages by 0.9%. The corresponding standard error equals 0.002, and thus the immediate wage effect is statistically highly significant. Moreover, the middle panel of the table shows that there is substantial persistence of this negative wage effect, and a negative and significant effect of initial labor market conditions remains as much as twenty years after first entry into the labor market. In fact, the point estimate of the semi-elasticity at twenty years of experience is even larger than the immediate wage penalty.

We add our small set of individual-level controls (i.e. our proxy for schooling and two dummy variables indicating a worker's initial occupation) in the second specification. This has some impact on the estimated wage effects, especially in the year of entry into the labor market (the immediate semi-elasticity of wages decreases from 0.009 to 0.007), but otherwise the coefficients are remarkably similar and a negative and strongly persistent wage effect of bad entry conditions remains, as the middle panel of table 2.2 reflects.

¹⁶We also computed standard errors that simultaneously account for clustering at both levels for our main estimates. This yields standard errors that are virtually indistinguishable from those actually reported.

Table 2.2: The long-run wage effects of initial labor market conditions

| Dependent variable | ln(real daily wage) | | | | | |
|------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Mean | 4.257 | 4.257 | 4.257 | 4.257 | 4.257 | 4.257 |
| Standard deviation | 0.380 | 0.380 | 0.380 | 0.380 | 0.380 | 0.380 |
| ur^0 | -0.009*** (0.002) | -0.007*** (0.002) | -0.009*** (0.002) | -0.009*** (0.002) | -0.010*** (0.002) | -0.009*** (0.002) |
| $exp \cdot ur^0$ | 0.000 (0.000) | 0.000 (0.000) | 0.001** (0.000) | 0.001** (0.000) | 0.000 (0.000) | 0.000 (0.000) |
| $exp^2 \cdot ur^0$ | -0.000*** (0.000) | -0.000*** (0.000) | -0.000*** (0.000) | -0.000*** (0.000) | -0.000* (0.000) | -0.000* (0.000) |
| $exp^3 \cdot ur^0$ | 0.000*** (0.000) | 0.000*** (0.000) | 0.000*** (0.000) | 0.000*** (0.000) | 0.000* (0.000) | 0.000* (0.000) |
| $\varepsilon_{ur}^y(0)$ | -0.009*** (0.002) | -0.007*** (0.002) | -0.009*** (0.002) | -0.009*** (0.002) | -0.010*** (0.002) | -0.009*** (0.002) |
| $\varepsilon_{ur}^y(5)$ | -0.011*** (0.002) | -0.010*** (0.002) | -0.010*** (0.002) | -0.010*** (0.002) | -0.010*** (0.001) | -0.009*** (0.001) |
| $\varepsilon_{ur}^y(10)$ | -0.017*** (0.002) | -0.016*** (0.002) | -0.015*** (0.002) | -0.015*** (0.001) | -0.012*** (0.001) | -0.012*** (0.001) |
| $\varepsilon_{ur}^y(15)$ | -0.021*** (0.002) | -0.020*** (0.002) | -0.019*** (0.002) | -0.019*** (0.002) | -0.015*** (0.002) | -0.014*** (0.001) |
| $\varepsilon_{ur}^y(20)$ | -0.015*** (0.003) | -0.015*** (0.002) | -0.014*** (0.003) | -0.014*** (0.003) | -0.018*** (0.002) | -0.017*** (0.002) |
| Potential experience | Yes | Yes | Yes | Yes | Yes | Yes |
| Longitudinal controls | Yes | Yes | No | Yes | Yes | Yes |
| Year of entry | Yes | Yes | Yes | Yes | Yes | Yes |
| Region of entry | Yes | Yes | Yes | Yes | Yes | Yes |
| Individual controls | No | Yes | Yes | Yes | Yes | Yes |
| Time dummies (biannual) | No | No | Yes | Yes | No | No |
| Year of entry $\times \kappa(exp)$ | No | No | No | No | Yes | Yes |
| Number of entrants | No | No | No | No | No | Yes |
| Number of observations | 3,349,075 | 3,349,075 | 3,349,075 | 3,349,075 | 3,349,075 | 3,349,075 |
| Adjusted R-Squared | 0.286 | 0.325 | 0.326 | 0.326 | 0.327 | 0.327 |
| Lifetime loss | 0.016 | 0.015 | 0.015 | 0.014 | 0.013 | 0.013 |

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by state at entry \times year of entry. exp and ur^0 denote potential labor market experience (in years) and the initial unemployment rate (in percentages), respectively. $\varepsilon_{ur}^y(k)$ denotes the estimated semi-elasticity of wages with respect to the initial unemployment rate, evaluated at k years of potential labor market experience.

We replace our set of longitudinal controls with a set of biannual year dummies in the third column. This yields estimates that are very similar to those from the specification with parametric business cycle controls. The fourth column includes both the set of biannual year dummies and our set of parametric longitudinal controls. Again, the estimates hardly change, although the estimated semi-elasticities at higher values of labor market experience are somewhat smaller than in the specification with only a parameterized trend.

We next add the full set of interaction terms between the year of entry and the

experience terms. This allows for changes in the wage profile across entry cohorts that are unrelated to initial conditions in a fully flexible way. The estimates again remain more or less stable.

We finally add the log number of labor market entrants aged 16-21 and 22-30 in the sixth column because there was a substantive drop in the relative number of labor market entrants aged 16-21 relative to all entrants aged below 30. The specification underlying the estimates shown in column 6 thus corresponds to the specification of equation (2.2). Our baseline specification still yields an immediate negative and statistically significant wage effect of about -0.009. Moreover, we find the wage penalty resulting from entering the labor market in times of high local unemployment to be highly persistent. In fact, the wage penalty is even slightly increasing with labor market experience, similar to what is found in Stevens (2007).

2.3.4 Robustness

Table 2.3 presents some different robustness checks. For the ease of comparison, the first column simply reproduces our baseline estimates from column 6 of table 2.2.

A first robustness check relates to the fact that a non-negligible number of wage observations is censored.¹⁷ Column 2 thus shows quantile (i.e. median) regression estimates of our baseline specification. The short-run estimates are identical but the estimated wage losses turn out to be moderately larger at higher values of labor market experience. More specifically, $\varepsilon_{ur}^0(5)$ is already slightly larger when estimated using a quantile regression, but the quantile regression estimate of $\varepsilon_{ur}^0(20)$ is as much as 35% larger than the corresponding OLS estimate. If anything, censored wages will thus lead us to underestimate the wage effects resulting from labor market entry in times of high local unemployment.

We add the interaction terms between our longitudinal controls and the experience terms in the third column to allow for the possibility that variations in the business cycle affect workers differentially at different stages of their labor market career. This, however, has no discernible effect on the estimated coefficients of our key parameters. Most of the estimated semi-elasticities are virtually identical to our baseline estimates.

We use different unemployment rates in columns 4 and 5 as another robustness check. We use unemployment rates at the NUTS-3 rather than at the state level in the fourth column (cf. footnote 9). Using the unemployment rate at the NUTS-3 level results in a smaller initial and a somewhat less persistent wage effect. This

¹⁷Appendix table A.1 shows that only few wage observations are censored in the year of entry. However, top censored wages become much more frequent as workers accumulate labor market experience.

is most likely due to the fact that we also include entry-region fixed effects at a finer aggregation level which tends to pick up more of the variation in wages.¹⁸ The specification in the fifth column uses state-level youth unemployment rates instead of overall unemployment rates. The resulting wage effects are somewhat larger, but still reasonably close to our main estimates based on overall unemployment rates.

Table 2.3: Sensitivity analysis

| Dependent variable | ln(real daily wage) | | | | | ln(earn.) |
|---------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Mean | 4.257 | 4.257 | 4.257 | 4.257 | 4.257 | 9.991 |
| Standard deviation | 0.380 | 0.380 | 0.380 | 0.380 | 0.380 | 0.656 |
| ur^0 | -0.009*** (0.002) | -0.009*** (0.000) | -0.009*** (0.002) | -0.005*** (0.001) | -0.006*** (0.002) | -0.007*** (0.003) |
| $exp \cdot ur^0$ | 0.000 (0.000) | -0.000 (0.000) | 0.000 (0.000) | -0.000 (0.000) | -0.001** (0.000) | -0.001* (0.001) |
| $exp^2 \cdot ur^0$ | -0.000* (0.000) | -0.000** (0.000) | -0.000* (0.000) | -0.000** (0.000) | 0.000 (0.000) | 0.000 (0.000) |
| $exp^3 \cdot ur^0$ | 0.000* (0.000) | 0.000 (0.000) | 0.000** (0.000) | 0.000*** (0.000) | 0.000 (0.000) | 0.000 (0.000) |
| $\varepsilon_{ur}^y(0)$ | -0.009*** (0.002) | -0.009*** (0.000) | -0.009*** (0.002) | -0.005*** (0.001) | -0.006*** (0.002) | -0.007*** (0.003) |
| $\varepsilon_{ur}^y(5)$ | -0.009*** (0.001) | -0.011*** (0.000) | -0.009*** (0.001) | -0.007*** (0.001) | -0.011*** (0.002) | -0.012*** (0.002) |
| $\varepsilon_{ur}^y(10)$ | -0.012*** (0.001) | -0.015*** (0.000) | -0.012*** (0.001) | -0.010*** (0.001) | -0.014*** (0.002) | -0.016*** (0.002) |
| $\varepsilon_{ur}^y(15)$ | -0.014*** (0.001) | -0.019*** (0.000) | -0.015*** (0.001) | -0.013*** (0.001) | -0.016*** (0.002) | -0.019*** (0.002) |
| $\varepsilon_{ur}^y(20)$ | -0.017*** (0.002) | -0.023*** (0.000) | -0.017*** (0.002) | -0.015*** (0.001) | -0.017*** (0.002) | -0.021*** (0.003) |
| Estimation method | OLS | LAD | OLS | OLS | OLS | OLS |
| Unemployment rate | State | State | State | NUTS-3 | Youth | State |
| Baseline controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Longit. controls $\times \kappa(exp)$ | No | No | Yes | No | No | No |
| Number of observations | 3,349,075 | 3,349,075 | 3,349,075 | 3,349,075 | 3,349,075 | 3,349,075 |
| Adjusted R-Squared | 0.327 | | 0.328 | 0.330 | 0.327 | 0.223 |
| Lifetime loss | 0.013 | 0.016 | 0.013 | 0.010 | 0.014 | 0.016 |

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by state at entry \times year of entry. The specification in column 1 is the same as in column 6 of table 2.2. All specifications include the same set of controls as in column 6 of table 2.2 (“baseline controls”). The dependent variable is the log real daily wage in columns 1-5 and log real annual earnings in column 6. We use the state unemployment rate for men aged 16-65 (“State”) in columns 1-3 and 6, the NUTS-3 level unemployment rate for men aged 16-65 (“NUTS-3”) in column 4, and the state-level unemployment rate for men aged 16-25 (“Youth”) in column 5.

¹⁸It may also reflect endogeneity of the local unemployment rate at lower aggregation levels, due to the fact that workers move from regions with high unemployment to those with lower levels of unemployment (Wozniak, 2010).

Finally, the specification in the sixth column shows results when we use log annual earnings instead of log daily wages as the dependent variable. In this case we get somewhat larger, but still very similar estimates than in the baseline case. The comparison with the effect on wages also implies that workers entering during times of high local unemployment not only suffer from lower subsequent wages, but also from less employment days throughout their labor market career (see also section 2.5 below where we discuss the effect of initial conditions on job mobility and unemployment).

2.3.5 Endogenous Labor Market Entry

Instrumental Variable Estimates

One important issue that we have not yet considered is the fact that the composition of the sample of labor market entrants may be endogenous with respect to variation in the local business cycle, which may lead to inconsistent estimates of the effect of initial conditions on wages.¹⁹ Various studies have tried to tackle this issue by instrumenting the unemployment rate at labor market entry with the unemployment rate at some earlier age or with the unemployment rate at the predicted date of graduation.²⁰ We have thus tried a similar approach, instrumenting the local unemployment rate at entry with the local unemployment rate at age 16, the age at which mandatory schooling ends in Austria.

However, we have some serious reservations, at least in our context, regarding the validity of this instrumental variable strategy. Most importantly, the instrument may have a direct effect on subsequent wages, in which case IV estimates will be biased (e.g. Angrist and Krueger, 1994). To see why this may occur, note that, the majority of individuals, at least in our sample, start an apprenticeship training at the age of 16, as this is the age when mandatory schooling ends. In Austria, moreover,

¹⁹Bils (1985) and more recently Solon *et al.* (1994) and Blundell *et al.* (2003) have put forth this line of argument. In fact, the timing of labor force entry and thus the composition of labor market cohorts may be endogenous for several distinct reasons. First, some potential labor market entrants may refrain from entering the labor market altogether. Second, both the choice of education as well as the duration of schooling may be endogenous, as both job prospects are weak and opportunity costs of schooling low in times of high unemployment. Third, some workers may simply delay their entry when faced with unfavorable entry conditions, either by registering for unemployment benefits or by staying out of the labor force until they find a job. Whatever the underlying reason, if those workers who do not immediately get a job are a selected group of all workers who intend to enter employment in a given year, then the composition of the actual entrants changes along with corresponding changes in the unemployment rate and thus potentially biases the estimated effect of the initial unemployment rate on wages.

²⁰Kahn (2010), Kondo (2008) and Oreopoulos *et al.* (2012) use a similar instrumental variable strategy. OLS and IV estimates are similar in Oreopoulos *et al.* (2012), but IV are substantially larger in both Kondo (2008) and Kahn (2010). Kwon *et al.* (2010) and Stevens (2007), the two European studies, only show OLS estimates.

most of the practical training is directly provided by firms which are free to decide whether they want to offer apprenticeships. It is thus quite obvious that labor market conditions may have an impact on firms' decision to offer apprenticeships.²¹ Moreover, economic conditions may not only impact the number, but also the type (i.e. quality) of apprenticeships firms offer.

Table 2.4 shows OLS and instrumental variable estimates. As before, the first column replicates our baseline estimates from column 6 of table 2.2. We first compare our baseline estimates to the corresponding 2SLS estimates that instrument the local unemployment rate at first entry with the local unemployment rate at age 16, shown in the fifth column. The 2SLS estimate of the semi-elasticity in the year of entry into the labor force is more than six times larger than the corresponding OLS estimate (-0.059 versus -0.009), and the lifetime loss in the case of 2SLS is about four times larger than in the baseline case (0.052 versus 0.013). While it appears reasonable that 2SLS estimates are larger than OLS estimates, signifying positive sample selection in economic downturns, the size of the difference between the estimates is almost suspiciously large. Thus let us examine the reduced form estimates as well, shown in column 3. Note that the immediate and medium-run wage losses are considerably larger than those in column 1, while the estimates appear to converge at higher values of labor market experience. It seems hard to rationalize the difference between columns 1 and 3 without considering the possibility of a direct effect of ur^{16} on the dependent variable.

Finally, since both our proxy for the completed duration of schooling and/or training as well as individuals' initial occupation may be endogenous with respect to the unemployment rate age 16, we also show estimates that do not control for these characteristics. Our baseline estimates are hardly any different, but note that the estimates using the unemployment rate at age 16 increase considerably, resulting in 2SLS estimates that are as much as ten times larger than the corresponding OLS estimates (compare columns 6 and 2). These estimates are presumably too large to be plausible, but we think they can be easily rationalized by acknowledging that the instrument has a direct impact on wages (through its impact on individuals' schooling/training) in which case 2SLS estimates are upwards biased.

²¹To the best of our knowledge, there is no relevant empirical evidence for Austria. However, Muehlemann *et al.* (2009) provide evidence consistent with this line of argument for Switzerland, which has an apprenticeship system very similar to that of Austria. Their results thus presumably carry over to Austria.

Table 2.4: OLS versus 2SLS estimates

| Dependent variable | ln(real daily wage) | | | | | |
|--------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Mean | 4.257 | 4.257 | 4.257 | 4.257 | 4.257 | 4.257 |
| Standard deviation | 0.380 | 0.380 | 0.380 | 0.380 | 0.380 | 0.380 |
| ur^0 | -0.009*** (0.002) | -0.011*** (0.002) | | | -0.059*** (0.006) | -0.135*** (0.017) |
| $exp \cdot ur^0$ | 0.000 (0.000) | 0.000 (0.000) | | | 0.005*** (0.001) | 0.005*** (0.001) |
| $exp^2 \cdot ur^0$ | -0.000* (0.000) | -0.000 (0.000) | | | -0.001*** (0.000) | -0.001*** (0.000) |
| $exp^3 \cdot ur^0$ | 0.000* (0.000) | 0.000 (0.000) | | | 0.000*** (0.000) | 0.000*** (0.000) |
| ur^{16} | | | -0.029*** (0.002) | -0.052*** (0.003) | | |
| $exp \cdot ur^{16}$ | | | 0.004*** (0.001) | 0.004*** (0.001) | | |
| $exp^2 \cdot ur^{16}$ | | | -0.000*** (0.000) | -0.000*** (0.000) | | |
| $exp^3 \cdot ur^{16}$ | | | 0.000*** (0.000) | 0.000*** (0.000) | | |
| $\varepsilon_{ur}^y(0)$ | -0.009*** (0.002) | -0.011*** (0.002) | -0.029*** (0.002) | -0.052*** (0.003) | -0.059*** (0.006) | -0.135*** (0.017) |
| $\varepsilon_{ur}^y(5)$ | -0.009*** (0.001) | -0.011*** (0.002) | -0.015*** (0.002) | -0.039*** (0.002) | -0.044*** (0.005) | -0.124*** (0.016) |
| $\varepsilon_{ur}^y(10)$ | -0.012*** (0.001) | -0.013*** (0.001) | -0.015*** (0.002) | -0.039*** (0.002) | -0.045*** (0.005) | -0.132*** (0.016) |
| $\varepsilon_{ur}^y(15)$ | -0.014*** (0.001) | -0.015*** (0.002) | -0.019*** (0.002) | -0.045*** (0.002) | -0.053*** (0.005) | -0.146*** (0.017) |
| $\varepsilon_{ur}^y(20)$ | -0.017*** (0.002) | -0.017*** (0.002) | -0.021*** (0.002) | -0.050*** (0.003) | -0.055*** (0.006) | -0.152*** (0.017) |
| Estimation method | OLS | OLS | OLS | OLS | 2SLS | 2SLS |
| Baseline controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Individual controls | Yes | No | Yes | No | Yes | No |
| Number of observations | 3,349,075 | 3,349,075 | 3,349,075 | 3,349,075 | 3,349,075 | 3,349,075 |
| Adjusted R-Squared | 0.327 | 0.287 | 0.328 | 0.296 | 0.323 | 0.249 |
| Lifetime loss | 0.013 | 0.014 | 0.019 | 0.046 | 0.052 | 0.143 |

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by state at entry \times year of entry. The first column replicates the baseline estimates from column 6 of table 2.2. All specifications include controls for potential labor market experience, year of entry, state at entry, longitudinal controls, the interactions between year of entry and $\kappa(exp)$, as well as controls for the number of labor market entrants (“baseline controls”). 2SLS estimates instrument the local unemployment rate at entry (ur^0) with the local unemployment rate at age 16 (ur^{16}).

Direct Evidence on Endogenous Labor Market Entry

We complement the evidence from our IV estimates with some more direct evidence on compositional effects from labor market entry conditions. A first piece of evidence is provided in table 2.5 that directly examines the association of entry conditions and the number of labor market entrants in different age brackets. In its most basic form, the estimated regression is:

$$\ln(e_{jt}) = ur_{jt}^0 \alpha + \psi_j + \theta_t + \varepsilon_{jt}, \quad (2.3)$$

with the dependent variable being the log number of labor market entrants in state j and year t . The unemployment rate is again denoted by ur_{jt}^0 . We also include state and year fixed effects, denoted by ψ_j and θ_t , respectively. The first panel of table 2.5 suggests that there is a non-negligible and negative effect of current labor market conditions on the number of entrants aged 16-21. The number of entrants aged 16-21 is estimated to decrease by 2.8% for an increase in the local initial unemployment rate by one percentage point. At the same time, the second and third column of table 2.5 show that there is no statistically significant effect on either the number of entrants aged 22-30 or the number of entrants aged 16-30, suggesting that some of the potential entrants with low or medium duration of schooling/training are permanently discouraged from entering the labor market at all, and thus that sample selection over the business cycle occurs through changes at the extensive margin.

Table 2.5: Number of labor market entrants

| Dependent variable | $\ln(\#entrants$ aged 16-21) | $\ln(\#entrants$ aged 22-30) | $\ln(\#entrants$ aged 16-30) |
|------------------------|---------------------------------|---------------------------------|---------------------------------|
| Mean | 7.782 | 7.058 | 8.229 |
| Standard deviation | 0.662 | 1.030 | 0.720 |
| ur^0 | -0.028*** (0.008) | 0.009 (0.012) | -0.001 (0.010) |
| State fixed effects | Yes | Yes | Yes |
| Time fixed effects | Yes | Yes | Yes |
| Number of observations | 207 | 207 | 207 |
| Adjusted R-Squared | 0.981 | 0.986 | 0.984 |

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses. The dependent variable is the log number of male labor market entrants in a given age group aggregated at state at entry \times year of entry cells.

Table 2.6 complements the evidence from table 2.5 by looking at the association between the initial local unemployment rate and our two proxies for schooling. Specifically, we regress our proxy for schooling s_i on the local initial unemployment rate and additional controls:

$$s_i = ur_{j[i]}^0 \alpha + \psi_j + \theta_t + \varepsilon_i, \quad (2.4)$$

with $ur_{j[i]}^0$ again denoting the local initial unemployment rate, and with ψ_j and θ_t denoting fixed effects for state at entry and year of entry, respectively. Based on the results from table 2.4 above, we also include the local unemployment rate at age 16 in some specifications. The results for our first measure of schooling, age at first entry into the labor market, yield somewhat mixed results (though the mixed results are probably due to the nature of our schooling proxies). While the first column shows a small positive impact of entry conditions, we get a non-significant estimate once we control for individuals' unemployment before their first regular employment. However, once we also include the unemployment rate at age 16, the coefficient on the local unemployment rate at entry turns significantly positive, while the point estimate on the unemployment rate at age 16 is statistically negative (consistent with our discussion of local labor market conditions at age 16 in section 2.3.5 above).

Table 2.6: Age at first entry into the labor market and age at start of first regular job

| Dependent variable | Age at labor market entry | | | | Age at start of first regular job | | |
|--------------------------|---------------------------|-------------------|-------------------|----------------------|-----------------------------------|----------------------|----------------------|
| | 18.669 | 18.669 | 18.669 | 18.669 | 19.195 | 19.195 | 19.195 |
| Mean | 1.431 | 1.431 | 1.431 | 1.431 | 1.007 | 1.007 | 1.007 |
| Standard deviation | | | | | | | |
| ur^0 | -0.037*** (0.013) | -0.010 (0.009) | -0.006 (0.009) | 0.438*** (0.032) | -0.007 (0.005) | 0.448*** (0.034) | 0.216*** (0.021) |
| ur^{16} | | | | -0.662*** (0.028) | | -0.679*** (0.029) | -0.328*** (0.022) |
| Age at first entry | | | | | | | 0.529*** (0.020) |
| State of entry | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year of entry | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Unempl. before first job | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of entrants | No | No | Yes | Yes | Yes | Yes | Yes |
| Number of observations | 217,587 | 217,587 | 217,587 | 217,587 | 217,587 | 217,587 | 217,587 |
| Adjusted R-Squared | 0.043 | 0.267 | 0.267 | 0.397 | 0.057 | 0.333 | 0.674 |

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered by state at entry \times year of entry.

The second part of table 2.6 shows similar results for our second proxy for schooling, age at the start of one's first regular job. Once we control for the unemployment rate at age 16, we get a positive coefficient on the local unemployment rate at entry.

Also note, however, that the estimated coefficients imply rather small effects on the average length of schooling/training in any case.²²

Taken together, the results from table 2.5 and table 2.6 suggest that there is some, arguably positive, sample selection in times of high local unemployment rates. For this reason, we believe that it is plausible to argue that our baseline estimates provide a lower bound on the true effect of labor market entry conditions on wages (or, rather, the effect of labor market conditions on wages in the absence of any sample selection issues). More importantly, the OLS estimates probably better reflect the overall effects of labor market entry conditions because the compositional effects, which also are caused by the cyclical variation in aggregate conditions, are factored in as well.

2.3.6 Quantitative Implications

Our estimates so far imply a non-negligible negative long-run wage effect from entering the labor market when the local unemployment rate is high. This is illustrated at the bottom of each table, which, where applicable, shows an approximate estimate of the lifetime loss in wages associated with a hypothetical increase in the initial unemployment rate by one percentage point. The lifetime loss in wages (or any other outcome considered) is computed simply as the average of the accumulated wage losses within the first twenty years of labor market experience.

Our estimates from tables 2.2 and 2.3 imply that an increase in the initial local unemployment rate by one percentage point is associated with an approximate lifetime loss in wages in the range of about 1% to 1.6%, and our preferred specification yields an approximate lifetime loss of 1.3% for an increase in the local initial unemployment rate by one percentage point. Moreover, according to our argumentation from the preceding section, this likely represents a lower bound (which in turn is consistent with the fact that our 2SLS estimates are larger than our baseline estimates).

Finally, to better understand the true size and economic meaning of these numbers, note that the estimated returns to apprenticeship training in Austria are on the order of about 3% to 5% per year (Fersterer *et al.*, 2008). The estimated wage loss resulting from an increase in the local initial unemployment rate by one percentage point thus approximately equals the value of about 0.26-0.4 a year of apprenticeship training.²³

²²This result is perhaps not surprising because apprenticeship positions usually have a fixed duration of either three or four years. Thus, in principle, it may be possible to observe no change in the average length of schooling/training even if there is considerable downranking to lower paying jobs.

²³It is reassuring that the estimated coefficient on the proxy for schooling/training equals about 0.046 in our baseline specification (i.e. column 6 of table 2.2), which is consistent with the reported external evidence on the returns on apprenticeship training in Austria.

This comparison emphasizes that unfavorable labor market entry conditions indeed have a sizable negative effect on workers' lifetime wages.

2.4 Initial Job Assignment

2.4.1 Quality of Workers' First Employer

Previous evidence strongly suggests that a worker's first employer may play an important role in explaining not only the immediate, but also the longer-run wage effects resulting from labor market entry conditions (Oreopoulos *et al.*, 2012; von Wachter and Bender, 2006). Different mechanisms may explain the observed importance of individuals' initial placement. One important mechanism that may explain cohort effects in wages starts with the observation that there is cyclical variation in job and/or task assignment within jobs over the business cycle and that high ability workers are assigned to jobs with lower average quality in recessions (e.g. Devereux, 2000).²⁴ If, moreover, these jobs offer different opportunities for accruing human capital or if, alternatively, a substantial part of human capital accumulated on the job is specific to a worker's task, occupation, and/or industry, then initial job and/or task assignment is probably able to explain a significant part of the observed persistent wage effects resulting from unfavorable initial labor market conditions (e.g. Gibbons and Waldman, 2006).²⁵ A similar argument is put forth in a more recent study by Liu *et al.* (2012), who argue that cyclical skill mismatch can generate persistent wage effects if at least part of the human capital acquired in a worker's first job/task is non-transferable to subsequent jobs/tasks and assuming that previously acquired (and potentially superior) human capital depreciates while not put to use.

Inspired by this evidence, we also provide an analysis of the association of initial labor market conditions and the quality of a worker's first job, which we complement with evidence on search frictions. We use a variety of different measures of the quality of workers' first employment. The first two measures are firm-level compensation measures, but we also look at a firm's size and age and the possibility that it might "disappear" within the following year or the following four years.²⁶ We regress each

²⁴In line with these findings, Mansour (2009) finds that the quality of a worker's first job is lower when entry happens during high unemployment, even though cohorts entering during recession are positively selected.

²⁵See Gathmann and Schönberg (2010) for evidence on task specific, Kambourov and Manovskii (2009) on occupation specific, and Neal (1995) or Parent (2000) on industry specific human capital. Sullivan (2010) provides evidence that both occupation and industry specific human capital are decisive in determining the level of wages.

²⁶The "disappearance" of a firm identifier does not necessarily coincide with a plant closure because

of these measures on the initial local unemployment rate and a couple of additional control variables:

$$\omega_i^{k(t_0)} = ur_{j[i]}^0 \alpha + \psi_i^{j(t_0)} + \phi_i^{t_0} + x_i^0 \beta_1 + \ln(n_{j[i]}^0) \beta_2 + \epsilon_i, \quad \text{or} \quad (2.5a)$$

$$\omega_i^{k(t_0)} = ur_{j[i]}^0 \alpha + \psi_i^{j(t_0)} + \phi_i^{t_0} + x_i^0 \beta_1 + \ln(n_{j[i]}^0) \beta_2 + \lambda_i^{k(t_0)} + \epsilon_i, \quad (2.5b)$$

where $\omega_i^{k(t_0)}$ denotes some outcome for individual i who first enters the labor market in year t_0 at employer k . The local initial unemployment rate $ur_{j[i]}^0$ is again the regressor of main interest. We also include the same set of individual-level controls x_i^0 , the log number of labor market entrants $n_{j[i]}^0$, as well as a full set of dummy variables indicating the state a worker resides in at entry, $\psi_i^{j(t_0)}$, and year of entry, $\phi_i^{t_0}$. As above, standard errors are clustered by year at entry \times state at entry. We show results with and without controls for an employer's industry affiliation $\lambda_i^{k(t_0)}$ because of potential endogeneity. Estimates of α are shown in table 2.7; panel (a) shows estimates without industry controls, panel (b) shows estimates including industry controls.

Consistent with previous evidence, we find considerable cyclical variation in most of our performance measures for workers' initial employer. Not surprisingly, the first two columns of table 2.7 show that entry in times of high local unemployment is associated with starting one's labor market career with an employer who pays considerably lower compensation on average. Initial employers of workers entering at different labor market conditions also differ in other dimensions, however. Specifically, columns 3-6 show that workers entering during times of high local unemployment also tend to enter the labor market through firms that are younger and smaller than first employers of workers who enter the labor force when the labor market is tight. More remarkably, the final two columns show that those individuals entering when unemployment is high are also much more likely to end up with an employer who is doomed to fail in the near future.

The comparison between the two panels of table 2.7 further suggests that some of the downgrading during bad entry conditions occurs both across and within industries. For example, the first two columns show that the estimated coefficients on wages and earnings are smaller if we control for industry affiliation of workers' initial employers, implying that part of the negative effect on wages/earnings stems from the fact that workers are more likely to start working with an employer from an industry that offers lower compensation.

a firm identifier may also disappear for other reasons (e.g. because of a take-over by another firm).

Table 2.7: Quality of first employer

| Dependent variable | Log mean compensation | | Firm age | | Firm size | | | Firm identifier disappears | |
|--|-----------------------|----------------------|----------------------|--|--------------------|-------------------|----------------------|----------------------------|---------------------|
| | Wages | Earnings | | | Small | Medium | Large | within 1 year | within 4 years |
| Mean | 4.022 | 9.638 | 10.392 | | 0.415 | 0.367 | 0.218 | 0.032 | 0.118 |
| Standard deviation | 0.310 | 0.445 | 6.619 | | 0.493 | 0.482 | 0.413 | 0.176 | 0.322 |
| <i>(a) Without controls for industry affiliation</i> | | | | | | | | | |
| ur^0 | -0.016*** (0.002) | -0.026*** (0.002) | -0.124*** (0.033) | | 0.005** (0.002) | 0.003 (0.002) | -0.007*** (0.002) | 0.004*** (0.001) | 0.013*** (0.002) |
| Adjusted R-Squared | 0.144 | 0.077 | 0.337 | | 0.030 | 0.007 | 0.050 | 0.005 | 0.011 |
| <i>(b) With controls for industry affiliation</i> | | | | | | | | | |
| ur^0 | -0.011*** (0.001) | -0.015*** (0.002) | -0.026 (0.034) | | 0.003 (0.002) | 0.003* (0.002) | -0.007*** (0.002) | 0.008*** (0.001) | 0.023*** (0.003) |
| Adjusted R-Squared | 0.447 | 0.389 | 0.407 | | 0.193 | 0.071 | 0.296 | 0.067 | 0.183 |
| Individual controls | Yes | Yes | Yes | | Yes | Yes | Yes | Yes | Yes |
| State of entry | Yes | Yes | Yes | | Yes | Yes | Yes | Yes | Yes |
| Year of entry | Yes | Yes | Yes | | Yes | Yes | Yes | Yes | Yes |
| Number of entrants | Yes | Yes | Yes | | Yes | Yes | Yes | Yes | Yes |
| Number of observations | 217,587 | 217,587 | 217,587 | | 217,587 | 217,587 | 217,587 | 217,587 | 217,587 |

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered on state at entry \times year of entry. Small (medium, large) firms are firms with 1-25 (26-250, 251+) employees.

Because we neither have direct nor detailed measures of the schooling and training actually acquired, we cannot provide any direct assessment on the importance of skill mismatch as in Liu *et al.* (2012). We can, however, provide some direct evidence on the cyclical nature of search frictions (which seems to be a natural precondition for skill mismatch). Table 2.8 shows that workers who enter during high unemployment not only end up working at a firm that is economically weaker, but that they also have a more difficult time of finding an employer at all. Specifically, the first three columns show that a higher local initial unemployment rate is associated with an increased likelihood of experiencing an unemployment spell before starting regular employment, with longer unemployment, as well as an increased probability of experiencing long-term unemployment even before starting one's first regular employment spell. Finally, the fourth column shows that workers entering during an economic downturn are also more likely to start their labor market career through marginal (i.e. poorly paid) employment.

Table 2.8: Search frictions

| Dependent variable | Unemployment before first regular job | | | Low starting wage |
|------------------------|---------------------------------------|---------------------|---------------------|---------------------|
| | Incidence | Duration | Longterm | |
| Mean | 0.306 | 42.649 | 0.018 | 0.025 |
| Standard deviation | 0.461 | 101.181 | 0.135 | 0.156 |
| ur^0 | 0.009*** (0.002) | 3.698*** (0.379) | 0.005*** (0.001) | 0.004*** (0.001) |
| Individual controls | Yes | Yes | Yes | Yes |
| State of entry | Yes | Yes | Yes | Yes |
| Year of entry | Yes | Yes | Yes | Yes |
| Number of entrants | Yes | Yes | Yes | Yes |
| Number of observations | 217,587 | 217,587 | 217,587 | 217,587 |
| Adjusted R-Squared | 0.305 | 0.316 | 0.083 | 0.012 |

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered on state at entry \times year of entry. Longterm unemployment before first regular job denotes unemployment of more than one year. Low starting wage denotes that the entry wage is below 1.2 times the lower censoring point (cf. footnote 8).

2.4.2 The Importance of the Initial Employer for the Evolution of Wages

Given this pronounced effects of initial conditions on various quality measures of workers' first employer, we next focus on the importance of workers' first employer on the evolution of their starting as well as subsequent wages. To this end we simply augment our baseline model, i.e. equation (2.2), with additional controls for a worker's first employer in a first step.

The first column of table 2.9 again replicates our baseline estimates, without any controls for an individual's initial employer. In contrast, the specification in the second column shows estimates from a regression that includes fixed effects for a worker's initial employer and thus controls for all time-invariant employer characteristics (whether observed or not).²⁷ Estimates turn out to be markedly different from those of our baseline specification. First, the estimated wage effect in the year of entry is much smaller than in the baseline specification (-0.002 compared to -0.009) and, moreover, the estimate is no longer statistically significant. Thus controlling for workers' initial employer seems to pick up not only all of the short-run but also most of the long-run wage effect from poor entry conditions. Indeed, while the life-time loss in wages amounts to about 1.3% in our baseline specification, controlling for both observed and unobserved characteristics of first employer only yields an estimate of about 0.4%. Thus, similar to Oreopoulos *et al.* (2012), controlling for the characteristics of workers' first employer explains most of the observed wage effects.

To explore the remaining negative wage effect in the longer run in somewhat more detail, we next split the sample by initial occupation (i.e. blue-collar versus white-collar occupations).²⁸ Columns 4 and 6 show results for blue-collar and white-collar workers separately without controls for the initial employer, while columns 5 and 7 show results with these controls. This shows that there are remarkably distinct results by initial occupation. Specifically, even though the initial wage effect is statistically similar for blue and white-collar workers, the longer-run effect differs substantially between the two groups. Specifically, because there is much less persistence in wage losses from poor entry conditions, the long-run wage effects for white collar workers are only about one-quarter the size of those for blue-collar workers (our estimates imply an approximate loss in lifetime wages of only 0.3% for white-collar workers,

²⁷We further include controls for time-varying characteristics of initial employers in column 3, but these variables do not have much additional impact on the estimated wage effects on top of the fixed effects for first employer.

²⁸Note that these two groups do not exactly add up to the overall sample size because some employment spells cannot be uniquely identified as either blue or white-collar (cf. table 2.1).

Table 2.9: Controlling for workers' first employer

| Dependent variable | ln(real daily wage) | | | | | | |
|----------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|---------------------|
| Mean | 4.257 | 4.257 | 4.257 | 4.241 | 4.241 | 4.367 | 4.367 |
| Standard deviation | 0.380 | 0.380 | 0.380 | 0.354 | 0.354 | 0.438 | 0.438 |
| ur^0 | -0.009*** (0.002) | -0.002 (0.002) | -0.002 (0.002) | -0.007*** (0.002) | 0.002 (0.001) | -0.010*** (0.002) | -0.002 (0.002) |
| $exp \cdot ur^0$ | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | -0.000 (0.000) | -0.000 (0.000) | 0.002*** (0.001) | 0.002*** (0.000) |
| $exp^2 \cdot ur^0$ | -0.000* (0.000) | -0.000 (0.000) | -0.000 (0.000) | -0.000 (0.000) | -0.000 (0.000) | -0.000*** (0.000) | -0.000** (0.000) |
| $exp^3 \cdot ur^0$ | 0.000* (0.000) | 0.000 (0.000) | 0.000 (0.000) | 0.000* (0.000) | 0.000 (0.000) | 0.000** (0.000) | 0.000* (0.000) |
| $\varepsilon_{ur}^y(0)$ | -0.009*** (0.002) | -0.002 (0.002) | -0.002 (0.002) | -0.007*** (0.002) | 0.002 (0.001) | -0.010*** (0.002) | -0.002 (0.002) |
| $\varepsilon_{ur}^y(5)$ | -0.009*** (0.001) | -0.003* (0.001) | -0.002* (0.001) | -0.009*** (0.002) | -0.001 (0.001) | -0.004** (0.002) | 0.005*** (0.002) |
| $\varepsilon_{ur}^y(10)$ | -0.012*** (0.001) | -0.004*** (0.001) | -0.004*** (0.001) | -0.013*** (0.002) | -0.004*** (0.001) | -0.002 (0.002) | 0.008*** (0.002) |
| $\varepsilon_{ur}^y(15)$ | -0.014*** (0.001) | -0.007*** (0.001) | -0.006*** (0.001) | -0.015*** (0.002) | -0.006*** (0.002) | -0.002 (0.003) | 0.009*** (0.002) |
| $\varepsilon_{ur}^y(20)$ | -0.017*** (0.002) | -0.008*** (0.002) | -0.007*** (0.002) | -0.016*** (0.002) | -0.007*** (0.002) | 0.000 (0.003) | 0.011*** (0.003) |
| Baseline controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm fixed-effects | No | Yes | Yes | No | Yes | No | Yes |
| Time-varying firm controls | No | No | Yes | No | Yes | No | Yes |
| BC ⁰ only | No | No | No | Yes | Yes | No | No |
| WC ⁰ only | No | No | No | No | No | Yes | Yes |
| Number of observations | 3,349,075 | | | 1,797,908 | | 563,764 | |
| Adjusted R-Squared | 0.327 | 0.504 | 0.507 | 0.247 | 0.473 | 0.437 | 0.635 |
| Lifetime loss | 0.013 | 0.005 | 0.004 | 0.013 | 0.004 | 0.003 | -0.007 |

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered on state at entry \times year of entry. The first column replicates the baseline estimates from column 6 of table 2.2.

compared to about 1.3% for blue-collar workers).²⁹ This result may emphasize that initial job and/or task assignment and occupation specific human capital are important in explaining persistent wage effects especially because occupation specific human capital is arguably more important for blue (e.g. craftsmen) than for white-collar jobs (e.g. Sullivan, 2010). We provide additional evidence that appears to be consistent and complementary with this explanation in the following section.

²⁹This result is somewhat surprising because previous research has shown that white-collar workers in Austria suffer from much larger and more persistent wage losses from job displacement (Schwerdt *et al.*, 2010). However, we use a totally different sample of workers than Schwerdt *et al.* (2010). Moreover, our study focuses on low and medium-skilled labor market entrants who entered the labor force between 1978 and 2000, while Schwerdt *et al.* (2010) focus on prime age workers of any educational level who experienced a plant closure between 1982 and 1988.

2.5 Job Mobility and Quality of Subsequent Employers

In the final part of the empirical analysis we focus on workers' mobility behavior and study how the quality of workers' employer changes as their labor market career proceeds, complementing the evidence from the preceding section. Table 2.10 looks at workers' mobility across jobs and, because some of this mobility may be frictional (if it is caused, for example, by the bankruptcy of a worker's current employer), we also look at workers' annual unemployment experience. As in section 2.4.2 above, we separate the sample by initial occupation. Moreover, we exclude all observations with zero years of labor market experience so that we can focus more clearly on workers' labor market career following their first year of entry.

The first two columns of table 2.10 show that blue-collar workers are somewhat more mobile, especially in the first years of their labor market career, if they entered the labor market when local unemployment was high. This makes sense, as we would expect workers to try to move to better employers as soon as possible if they had to start their labor market career with a low-performance employer. At the same time, however, columns 3 and 4 also show that these workers face a substantively increased risk of experiencing an unemployment spell and, if they become unemployed, they suffer from substantively longer spells of unemployment. Thus blue-collar workers apparently suffer from a permanent increase in both the risk and duration of unemployment. Part of this increase in the risk and duration of unemployment obviously relates to the fact that blue-collar workers are more likely to start their labor market career with an unsuccessful employer (cf. table 2.7), another part can be related to corresponding persistent effects on the quality of subsequent employers, as shown and discussed below.

We do not find any effect on job mobility for white-collar workers, as evident from column 5 and 6 (i.e. none of the relevant point estimates is statistically different from zero). We only find an increase in the duration of annual unemployment (without any increase in the incidence of unemployment, however), but one that is considerably smaller than the corresponding increase for blue-collar workers.

Table 2.10: Job mobility and unemployment

| Dependent variable | Job mobility | | Unemployment | | Job mobility | | Unemployment | |
|--------------------------|---------------------|-------------------|---------------------|---------------------|-------------------|-------------------|--------------------|--------------------|
| | New employer | New industry | Incidence | Duration | New employer | New industry | Incidence | Duration |
| Mean | 0.219 | 0.167 | 0.267 | 23.250 | 0.160 | 0.122 | 0.095 | 9.735 |
| Standard deviation | 0.413 | 0.373 | 0.443 | 52.060 | 0.367 | 0.327 | 0.293 | 39.269 |
| ur^0 | 0.002 (0.002) | 0.001 (0.002) | 0.008*** (0.002) | 1.090*** (0.269) | -0.001 (0.002) | -0.000 (0.001) | -0.001 (0.002) | -0.192 (0.227) |
| $exp \cdot ur^0$ | 0.000 (0.001) | -0.000 (0.000) | -0.001 (0.001) | 0.013 (0.074) | 0.001 (0.001) | 0.001 (0.001) | 0.000 (0.001) | 0.170** (0.081) |
| $exp^2 \cdot ur^0$ | -0.000 (0.000) | -0.000 (0.000) | 0.000 (0.000) | -0.005 (0.007) | -0.000 (0.000) | -0.000 (0.000) | -0.000 (0.000) | -0.015* (0.008) |
| $exp^3 \cdot ur^0$ | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) |
| $\varepsilon_{ur}^y(1)$ | 0.002* (0.001) | 0.001 (0.001) | 0.007*** (0.002) | 1.098*** (0.232) | -0.000 (0.001) | 0.000 (0.001) | -0.001 (0.001) | -0.036 (0.173) |
| $\varepsilon_{ur}^y(5)$ | 0.003*** (0.001) | 0.000 (0.001) | 0.005*** (0.002) | 1.051*** (0.203) | 0.001 (0.001) | 0.001 (0.001) | -0.000 (0.001) | 0.328** (0.139) |
| $\varepsilon_{ur}^y(10)$ | 0.002 (0.001) | -0.000 (0.001) | 0.003 (0.002) | 0.875*** (0.211) | -0.000 (0.001) | 0.001 (0.001) | -0.001 (0.001) | 0.345** (0.159) |
| $\varepsilon_{ur}^y(15)$ | 0.001 (0.001) | 0.000 (0.001) | 0.001 (0.002) | 0.662*** (0.225) | -0.002 (0.001) | -0.001 (0.001) | -0.002 (0.001) | 0.103 (0.189) |
| $\varepsilon_{ur}^y(20)$ | 0.002 (0.001) | 0.002* (0.001) | -0.000 (0.002) | 0.516** (0.239) | -0.003 (0.002) | -0.001 (0.002) | -0.003* (0.002) | -0.157 (0.226) |
| Baseline controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| BC ⁰ only | Yes | Yes | Yes | Yes | No | No | No | No |
| WC ⁰ only | No | No | No | No | Yes | Yes | Yes | Yes |
| Number of observations | 1,679,077 | 1,679,077 | 1,679,077 | 1,679,077 | 525,560 | 525,560 | 525,560 | 525,560 |
| Adjusted R-Squared | 0.026 | 0.021 | 0.045 | 0.029 | 0.017 | 0.015 | 0.025 | 0.018 |
| Lifetime loss | -0.002 | -0.000 | -0.003 | -0.836 | 0.001 | -0.000 | 0.001 | -0.172 |

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered on state at entry \times year of entry. Observations with zero years of labor market experience are excluded.

Table 2.11 finally shows how the quality of workers' subsequent employers changes over time (as in table 2.10, observations with zero years of experience are excluded). Not surprisingly, we find adjustment in compensation for both blue and white-collar workers but, consistent with our previous results from section 2.4.2 above, adjustment is apparently much slower and incomplete even after twenty years of experience for blue-collar workers. Apparently, adjustment for white-collar workers occurs much faster and appears to be complete ten years after first entry into the labor force at the latest. Moreover, wage adjustment for white-collar workers apparently occurs mainly within employers. Even more remarkably, we find that blue-collar workers are persistently exposed to an increased risk of job displacement, which has consistently been shown to be associated with substantive earnings losses (e.g. Jacobson *et al.*, 1993; Kletzer and Fairlie, 2003).

Table 2.11: Quality of Current Employer

| Dependent variable | Log mean firm wages | Log mean firm earnings | Firm id disappears wth.1 year | Log mean firm wages | Log mean firm earnings | Firm id disappears wth.1 year |
|--------------------------|---------------------------|------------------------------|-------------------------------------|---------------------------|------------------------------|-------------------------------------|
| Mean | 4.208 | 9.831 | 0.168 | 4.313 | 10.013 | 0.165 |
| Standard deviation | 0.288 | 0.459 | 0.374 | 0.296 | 0.422 | 0.371 |
| ur^0 | -0.010*** (0.002) | -0.020*** (0.003) | 0.004*** (0.001) | -0.009*** (0.002) | -0.015*** (0.003) | -0.000 (0.001) |
| $exp \cdot ur^0$ | 0.001*** (0.000) | 0.002*** (0.001) | -0.000 (0.000) | 0.002*** (0.001) | 0.002*** (0.001) | 0.000 (0.000) |
| $exp^2 \cdot ur^0$ | -0.000** (0.000) | -0.000*** (0.000) | 0.000 (0.000) | -0.000 (0.000) | -0.000 (0.000) | -0.000 (0.000) |
| $exp^3 \cdot ur^0$ | 0.000** (0.000) | 0.000*** (0.000) | -0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) |
| $\varepsilon_{ur}^y(1)$ | -0.009*** (0.002) | -0.018*** (0.003) | 0.003*** (0.001) | -0.008*** (0.002) | -0.013*** (0.002) | 0.000 (0.001) |
| $\varepsilon_{ur}^y(5)$ | -0.006*** (0.001) | -0.012*** (0.002) | 0.003*** (0.001) | -0.003* (0.002) | -0.005** (0.002) | 0.000 (0.001) |
| $\varepsilon_{ur}^y(10)$ | -0.005*** (0.001) | -0.009*** (0.002) | 0.003*** (0.001) | 0.001 (0.002) | -0.000 (0.002) | 0.000 (0.001) |
| $\varepsilon_{ur}^y(15)$ | -0.005*** (0.001) | -0.009*** (0.002) | 0.003*** (0.001) | 0.002 (0.002) | 0.002 (0.003) | -0.000 (0.001) |
| $\varepsilon_{ur}^y(20)$ | -0.004*** (0.002) | -0.009*** (0.002) | 0.003*** (0.001) | 0.003 (0.002) | 0.002 (0.003) | -0.001 (0.001) |
| Baseline controls | Yes | Yes | Yes | Yes | Yes | Yes |
| BC ⁰ only | Yes | Yes | Yes | No | No | No |
| WC ⁰ only | No | No | No | Yes | Yes | Yes |
| Number of obs. | 1,679,077 | 1,679,077 | 1,679,077 | 525,560 | 525,560 | 525,560 |
| Adjusted R ² | 0.124 | 0.084 | 0.575 | 0.167 | 0.108 | 0.610 |
| Lifetime loss | 0.006 | 0.010 | -0.003 | 0.000 | 0.002 | -0.000 |

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. Robust standard errors are given in parentheses and are clustered on state at entry \times year of entry.

Taken together, the fact that the adjustment in employer quality is incomplete for blue-collar workers, notwithstanding the increase in job mobility in at least the first five years of their labor market career, as well as the permanent increase in the risk of job displacement suggests that a share of the blue-collar workers entering the labor force in an economic downturn are forced into, and persistently locked in, low-paying jobs and/or tasks. Moreover, occupation and/or task-specific human capital may explain why these workers suffer permanent wages losses and why they are not able to move on to better jobs or tasks later on in their career.

2.6 Conclusions

We estimate the long-run impact of initial labor market conditions on wages for young males entering the Austrian labor market between 1978 and 2000. Consistent with previous evidence, we find a substantial wage penalty from poor entry conditions on starting wages. The estimated semi-elasticity of starting wages with respect to the initial unemployment rate is on the order of -0.005 to -0.010, with our preferred estimate equaling -0.009 (implying that a one percentage point increase in the initial local unemployment rate is associated with 0.9% lower starting wages). Moreover, this initial wage loss from first entry into the labor force during high local unemployment turns out to be highly persistent. Consequently, we find sizable negative effects on lifetime wages. Our preferred estimates imply that an increase in the initial local unemployment rate by one percentage point is associated with an approximate loss in lifetime wages of about 1.3%. Moreover, our estimates are close to the results reported in Stevens (2007) for Germany and Kwon *et al.* (2010) for Sweden. Our finding of persistent wage effects from poor entry conditions is also in line with Frühwirth-Schnatter *et al.* (2011) who study labor earnings dynamics in Austria using the same data source as we do, but a completely different empirical approach.

We further show that there is considerable cyclical variation in the quality of a worker's first regular employment. Even though workers are positively selected in times of high unemployment, they often end up working with employers that generally pay lower wages. We also show that initial job and/or task assignment within jobs is very important in explaining both the short and long-run wage effects from entry into the labor force during poor local entry conditions. In fact, controlling for the characteristics of workers' first employers can explain as much as three-quarters of the observed long-run wage effects resulting from poor entry conditions.

We finally find some very distinct patterns for blue and white-collar workers. The negative wage effects fade away for white-collar workers after about five to ten years,

while blue-collar workers suffer from very persistent negative wage effects from poor labor market entry conditions. At least part of this difference appears to be related to corresponding differences in the speed of quality adjustment of subsequent employers between occupations, where only blue-collar workers are permanently downgraded to lower quality employers.

2.A Appendix

Sample Construction

We first determine the start of the first regular employment spell for each male individual born between 1958 and 1985.³⁰ The restriction with respect to year of birth, in combination with the restriction on age at beginning of one's first regular job that we apply below, ensures that the potential range of age at first entry into the labor force is the same for each entry cohort considered in the analysis (1978–2000). Additionally, we drop all individuals who were self-employed and/or worked as a farmer or civil servant at least once, because the data do not consistently cover these employment spells over the whole period of analysis and/or because earnings are not recorded (in the case of self-employment). We thus cannot fully observe the employment and/or earnings histories of such individuals.

We then determine each individual's age at the start of his first regular employment spell starting between 1978 and 2000. We define regular employment as an employment spell which lasts for at least 180 days.³¹ We also focus on individuals aged between 16 and 21 years at the start of their first regular employment spell. This leaves us with 797,846 unique individuals (see table A.1), from which we take a simple 30% random sample. This finally yields a total of 217,587 unique individuals and 3,349,075 observations (= individuals \times years) when following these individuals over time.

Table A.1 shows descriptive statistics for some key variables, by individuals' age at the time of first entry into the labor force. We consistently exclude individuals who start their first regular employment after they attain age 30 because they presumably never enter the labor force at all.³² The first column of table A.1 shows descriptives

³⁰Obviously, an individual must be covered by the ASSD in order to be included in the sample. An individual is covered by the ASSD if he or she is entitled to future social security benefits (typically old-age benefits) or has already claimed social security benefits before first entering the labor force. Typically individuals "enter" the ASSD once they start working.

³¹Importantly, vocational training such as apprenticeship training is not considered as regular employment (but as formal training).

³²The group of individuals who enter at a later age probably consists of two very different groups,

Table A.1: Sample selection

| | Age at start of first regular employment | | | | | |
|-------------------------------|--|----------|---------|----------|---------|----------|
| | 16-30 | | 16-21 | | 22-30 | |
| Real daily wage | 50.488 | (18.644) | 47.536 | (15.303) | 56.742 | (23.027) |
| Age at start of first job | 20.863 | (2.948) | 19.200 | (1.004) | 24.386 | (2.589) |
| Duration of first regular job | 2.800 | (3.995) | 2.692 | (4.035) | 3.030 | (3.900) |
| Blue-collar | 0.650 | (0.477) | 0.726 | (0.446) | 0.487 | (0.500) |
| White-collar | 0.348 | (0.476) | 0.271 | (0.445) | 0.510 | (0.500) |
| Below GfGr: Yes = 1 | 0.059 | (0.235) | 0.050 | (0.219) | 0.077 | (0.266) |
| Above HBGr: Yes = 1 | 0.049 | (0.215) | 0.019 | (0.135) | 0.112 | (0.315) |
| <i>Region of employer:</i> | | | | | | |
| Vienna | 0.239 | (0.426) | 0.198 | (0.398) | 0.326 | (0.469) |
| Lower Austria | 0.158 | (0.365) | 0.171 | (0.377) | 0.129 | (0.336) |
| Burgenland | 0.024 | (0.152) | 0.000 | (0.000) | 0.021 | (0.142) |
| Upper Austria | 0.168 | (0.374) | 0.183 | (0.387) | 0.137 | (0.344) |
| Styria | 0.139 | (0.346) | 0.145 | (0.352) | 0.129 | (0.335) |
| Carinthia | 0.065 | (0.246) | 0.066 | (0.249) | 0.062 | (0.241) |
| Salzburg | 0.070 | (0.255) | 0.071 | (0.257) | 0.067 | (0.249) |
| Tyrol | 0.090 | (0.286) | 0.091 | (0.288) | 0.087 | (0.282) |
| Vorarlberg | 0.000 | (0.000) | 0.049 | (0.217) | 0.000 | (0.000) |
| <i>Industry of employer:</i> | | | | | | |
| Agriculture | 0.015 | (0.120) | 0.012 | (0.108) | 0.021 | (0.142) |
| Electricity | 0.007 | (0.085) | 0.009 | (0.095) | 0.003 | (0.057) |
| Mining | 0.006 | (0.078) | 0.007 | (0.082) | 0.005 | (0.068) |
| Manufacturing | 0.344 | (0.475) | 0.401 | (0.490) | 0.225 | (0.417) |
| Construction | 0.165 | (0.371) | 0.190 | (0.392) | 0.114 | (0.318) |
| Wholesale and retail trade | 0.150 | (0.357) | 0.159 | (0.366) | 0.130 | (0.336) |
| Gastronomy, hotel business | 0.060 | (0.237) | 0.046 | (0.209) | 0.089 | (0.285) |
| Transportation | 0.059 | (0.236) | 0.060 | (0.238) | 0.057 | (0.233) |
| Finance | 0.082 | (0.274) | 0.058 | (0.235) | 0.130 | (0.337) |
| Cleaning, body care | 0.009 | (0.096) | 0.008 | (0.089) | 0.012 | (0.109) |
| Arts, entertainment, sports | 0.010 | (0.102) | 0.005 | (0.072) | 0.021 | (0.145) |
| Healthcare, welfare | 0.014 | (0.116) | 0.006 | (0.079) | 0.029 | (0.167) |
| Education, research | 0.015 | (0.121) | 0.000 | (0.000) | 0.037 | (0.189) |
| Lobbies, social sec. agencies | 0.063 | (0.243) | 0.033 | (0.180) | 0.125 | (0.331) |
| Housekeeping | 0.000 | (0.000) | 0.000 | (0.022) | 0.000 | (0.000) |
| Number of observations | 1,174,523 | | 797,846 | | 376,677 | |

Notes: Table entries are sample means and standard deviations (in parentheses). Daily wages are given in Euros. Real wage are deflated using the consumer price index with base year 2007. GfGr (“Gerinfügigkeitsgrenze”) and HBGr (“Höchstbemessungsgrundlage”) denote the upper and the lower censoring point with respect to earnings, respectively. “Below GfGr” is an indicator taking on the value 1 if the nominal daily wage is equal to or below 1.2 times the lower censoring point and 0 otherwise. “Above HBGr” is an indicator taking on the value 1 if the nominal daily wage is equal to or above 0.8 times the upper censoring point and 0 otherwise.

for all individuals, the second (third) column shows descriptives for individuals aged 16 to 21 (aged 22 to 30) when starting their first employment. A comparison of the second to the third column shows that our sample restriction with respect to age at the start of first regular employment works as expected. The sample of lower skilled workers, compared to the group of higher skilled workers, contains a higher fraction of blue-collar workers, has considerably lower wages on average, and shorter duration of the first regular employment spell.

Also note that, for the group of individuals aged between 22 and 30 when first entering the labor force, highly skilled workers potentially are mixed up with low-skilled workers: individuals in this group of workers either spent much time in education, were previously unemployed, or had short employment spells not counting as regular employment. This is apparent from the proportion of workers below the lower censoring point or above the higher censoring point. The probability of crossing any of the two points is higher for the sample of older workers. Consequently, the variation in the real daily wage (and thus productivity) is considerably smaller in the sample of younger workers than in the group of older workers.

Acknowledgments

We thank Joshua Angrist, Johann Brunner, Pierre Cahuc, Christian Dustmann, Marcus Hagedorn, Martin Halla, Christian Hopenstrick, Helmut Hofer, Steinar Holden, Bo Honoré, Rafael Lalive, Michael Lechner, Andrew Oswald, Tamas Papp, Michael Reiter, Steven Stillman, Petra Todd, Till von Wachter, Rudolf Winter-Ebmer, Andrea Weber, Tobias Würzler, Josef Zweimüller, seminar participants in Linz, Vienna, Weggis and Zurich, as well as participants at the 2009 Engelberg Labor Economics Seminar, the 2009 Annual Meeting of the Austrian Economic Society in Linz, the 2009 Spring Meeting of Young Economists in Istanbul, the 15th International Conference on Panel Data in Bonn, and the 13th IZA European Summer School in Labor Economics in Buch/Ammersee for helpful comments and suggestions at various stages of this project. An earlier version of this paper appeared as “To Shape the Future: How Labor Market Entry Conditions Affect Individuals’ Long-Run Wage Profiles”. Financial support by the Austrian Science Fund is gratefully acknowledged (“The Austrian Center for Labor Economics and the Analysis of the Welfare State”).

who are indistinguishable from each other in the data. On the one hand there are truly high-skilled workers who enter the labor market at a later stage because they continued their education until that time. On the other hand, however, there are also low-skilled workers who were never employed or only sporadically employed before starting their first regular employment. Because schooling is not directly observed, we would mix these two groups of workers together if we were to include them.

CHAPTER 3

Announcement Effects of Health Policy Reforms: Evidence from the Abolition of Austria's Baby Bonus

Joint with Andreas Kuhn

A revised version of this chapter is accepted for publication in the European Journal of Health Economics. An earlier version has been published in 2011 as IZA Discussion Paper, No. 6141; ECON Working Paper, No. 48; NRN Working Paper, No. 1116.

“Economics is, at root, the study of incentives.”

Steven D. Levitt, American economist

3.1 Introduction

Forward looking behavior of individuals has potentially important implications for the implementation, as well as for the evaluation, of health policy reforms (and beyond). Indeed, if individuals are forward looking and if policy reforms are publicly announced in advance, individuals are likely to adapt their behavior even before the effective implementation of any given policy reform. As Malani and Reif (2010, p.2) put it, “anticipation is a reasonable diagnosis if individuals are forward looking, have access to information on future treatment, and there is a benefit to acting before the treatment is adopted”. Clearly, policy makers need to be aware of potential announcement

effects resulting from anticipatory behavior whenever they are planning to implement major health policy reforms and to announce them beforehand. It is also evident that anticipatory behavior should be factored in when evaluating specific health policy reforms. In fact, ignoring announcement effects may lead to misleading conclusions regarding the impact of the policy reform under study (Blundell *et al.*, 2011; Malani and Reif, 2010). Moreover, we may be especially concerned about potential negative health effects if individuals anticipate even small changes in financial incentives but are unable to foresee all potential health effects resulting from specific behavioral changes. We believe that such a situation is especially common in health policy contexts where it is very difficult, even for experts, to plausibly assess all potential health effects resulting from any given change in health-related behavior.

Some of the most compelling empirical evidence on announcement effects of health policy reforms available comes from the recent experience of introducing baby bonuses in Australia and Germany, respectively. Indeed, a couple of recent empirical studies has convincingly shown that the introduction of such policy measures is usually associated with considerable behavioral responses in the short-run, potentially resulting in negative health effects for the mother and/or her newborn child.¹ In the case of Australia, the government publicly announced on May 11, 2004, that it would pay 3000\$ (about 2190€) to each family of a newborn child born on or after July 1, 2004.² The introduction of the bonus was thus announced in advance of the effective policy change, creating an incentive to delay births. In their empirical evaluation of the policy change, Gans and Leigh (2009) find that 6% of the births (more than 1000 births) expected to happen in the 28 days preceding the actual policy change were moved to July 1st 2004 or later to become eligible for the baby bonus.³ They find that most of the effect is due to a corresponding timing of induction and cesarean section procedures. Consequently, they also find that children who were moved into the eligibility period were more likely to be of high birth weight.⁴ Tamm (2009) analyzes a similar

¹Substantial effects on birth timing are also found by Dickert-Conlin and Chandra (1999) and Chen (2011) who study the effects of tax incentives on the timing of births in the U.S. and France, respectively. Other studies have found that taxes distort other types of individual behavior such as marriages (Alm and Whittington, 1995) or even deaths (Kopczuk and Slemrod, 2003).

²More precisely, the baby bonus replaced another policy previously in force. In the previous system, the bonus was dependent on the income of the primary caregiver in the year the child was born and was in the form of a refundable tax offset. Most, though not all, households had an incentive to move births to July 1, 2004, or later. See Gans and Leigh (2009) for details.

³Drago *et al.* (2011) also analyze the introduction of the birth benefit in Australia, but use a different data source. They find that the birth benefit had both a positive effect on women's fertility intentions and one of modest size on the effective birthrate. Positive fertility effects from the Australian policy change are also reported in Lain *et al.* (2009).

⁴This in turn may imply long-run effects of short-run behavioral responses, since birth weight is suspected to be causally related with later labor-market outcomes (e.g. Black *et al.*, 2007).

reform in the system of family benefits in Germany, the introduction of parental leave benefits (“Elterngeld”) as of January 1, 2007.⁵ As in Australia, the announcement happened only a few weeks in advance, meaning that the policy change could only affect the timing, but not the number of births. Similarly to Gans and Leigh (2009), Tamm finds that a substantial number of births were delayed and moved into the eligibility period for the new benefit system. Specifically, he concludes that almost 8% of births (around 1000 births) that could have been expected in the last week of December were shifted to the first week of January 2007. He also finds a slight increase in birth weight among the births most likely to have been shifted (i.e. January versus December births).

In this chapter we study the fertility effects, as well as the potential health consequences for both mother and newborn child, following the announcement of the abolition of the Austrian baby bonus as of January 1, 1997. The Austrian baby bonus amounted to a maximum of 1090 € per child in 1996, the year before the abolition, and was paid conditional on medical examinations of both mother and newborn child. The unique feature of this policy change is that the elimination of the benefit was announced about ten months prior to enactment, creating the potential for an announcement effect because prospective parents had both an incentive and the opportunity to move their baby plans forward. Although the response window in order to qualify for the birth benefit before its abolition was only limited to three weeks, the early announcement could have increased the number of babies born in the month prior to the policy change. On top of this, pregnant women with a due date close to the date of the policy change might have manipulated the exact day of birth by means of a medical intervention (i.e. cesarean section). In the second part of the analysis we will explore whether the early announcement of the policy reform had any negative health effects for mothers and/or her newborn children.

The remainder of this chapter is organized as follows. In section 3.2 we provide some background information on the baby bonus in Austria. This is followed by a short discussion of the data source and some descriptives in section 3.3. We present our estimates of the fertility response following the announcement of the abolition of the baby bonus in section 3.4. Section 3.5 examines whether mothers (un)consciously take increased health risks for themselves and/or their babies when rescheduling the

⁵As in the case of Australia, the German policy changed incentives differently for households with different characteristics. Generally, households with women working before giving birth, those planning to work shortly following birth, or those with high income received higher benefits after the reform and thus had an incentive to delay their births. See Tamm (2009) for details. Neugart and Ohlsson (2011) provide an alternative evaluation of the German parental benefit reform (with similar conclusions).

timing of conception or birth. Section 4.6 concludes.

3.2 The Austrian Baby Bonus

3.2.1 Institutional Background

The Austrian baby bonus (“Geburtenbeihilfe”) was first introduced on January 1, 1968, as an untaxed single payment per live birth. In 1975 the payment of the bonus was made conditional on medical examinations both during pregnancy and after childbirth, and the payment of the bonus was consequently partitioned. After the last expansion of the birth benefit in January 1987, the maximum benefit amounted to 1090 € per child and was paid in five consecutive rates. The first rate of the bonus was paid immediately after birth (145.3 €), the second rate one week after birth (218 €), and the remaining three rates were paid after the child’s first (363.4 €), second (218 €), and fourth (145.3 €) birthday.

Policy makers wanted to sustain the incentive for mothers to continue with medical examinations for themselves and their newborn children even after the abolition of the baby bonus; they thus introduced an alternative incentive, the so-called “Mutter-Kind-Pass”, which is still in place today. It consists of a single bonus of 145 € per birth; the payment is conditional on both mother and child undergoing specific medical examinations, and it is paid on the child’s first birthday. Furthermore, eligibility to the new bonus is confined to mothers who are the child’s primary caregiver and to households whose income does not exceed a given threshold in the year of birth.⁶ Thus, depending on household income, the abolition of the birth benefit meant a cut in cash benefits amounting to either 945 € or 1090 € (equivalent to a cut in benefits of 87% or 100%, respectively). For a full-time employee (not household) with median labor earnings in 1996 equal to 20,991 €, the full amount of the baby bonus was worth approximately 4.5% of his or her annual earnings (equivalent to about 2.3 weeks’ income).

Compared to other family benefits, the baby bonus was rather modest in size. Aside from the birth benefit, three basic other types of family benefits existed (and still exist today). The most important (i.e. substantial) benefit is the family benefit (“Familienbeihilfe”), which is paid until the child’s 18th birthday at the minimum. In 1997 it amounted to 95-134 € per month, depending on the child’s age. Parental leave

⁶Specifically, the maximum household income in order to qualify for the “Mutter-Kind-Pass” bonus is defined as 11·HBGr, with HBGr (“Höchstbemessungsgrundlage”) denoting the upper income threshold above which the maximum pension benefit accrues. The threshold varies over time and amounted to 2965 € in 1997. Thus annual household income had to be lower than 32,616 € in 1997 to qualify for the “Mutter-Kind-Pass”.

benefits were paid over a period of 1.5 years at that time, and amounted to 340 € per month. However, until another major reform in 2002, these benefits were tied to the mother's employment before giving birth. Finally, there is a monthly tax allowance for children who live in the same household as the parent filing the tax report. The tax allowance amounted to 25-51 € per month in 1997, depending on the child's parity. Taken together, the baby bonus accounted for roughly 9% of all benefits (excluding tax allowances) accruing within the first four years of a child's life.⁷

3.2.2 The Abolition of the Baby Bonus

The structural deficit of the federal budget was the ultimate reason for the abolition of the baby bonus. Generous social benefits combined with a deterioration of the labor market caused the ratio of social expenditure to GDP to skyrocket in the early 1990s. In spite of a temporary strengthening of the economy in 1994, social expenditures still rose, resulting in an overall increase of 36.5% between 1991 to 1996 (Bauernberger and Guger, 1998). To decelerate rising social spending, the governing coalition between the conservatives and the social democrats finally passed an encompassing austerity package ("Strukturanpassungsgesetz") on July 1, 1996. Savings in family policy should be achieved by reducing maternity leave duration by half a year (from 24 to 18 months) and by abolishing the baby bonus. In terms of our identification strategy, it is important to stress that, except for the birth benefit, all reforms decided on within the framework of the austerity program came into effect on July 1, 1996 – half a year before the abolition of the baby bonus.

Our review of newspapers suggests that the abolition must have been known by February 2, 1996, when the coalition between the conservatives and the social democrats first announced their agreement on the austerity package. There was extensive press coverage, but there was also confusion about the exact date of abolition, and the media initially discussed July 1, 1996 as the effective date of elimination. By the first week of March, however, shortly before the coalition's agreement on the structural adjustment law was signed (March 11, 1996), it must have been evident that the birth benefit would be canceled for all children born on January 1, 1997 or later.

3.2.3 The Window of Opportunity

From what we have said above, it follows that there was a time gap of nearly ten months between the definitive announcement and the effective date of the policy

⁷Neglecting tax deductibles, $[1090 \text{ €} / (4 \cdot 12 \cdot 94.5 \text{ €} + 1.5 \cdot 12 \cdot 338.6 \text{ €} + 1090 \text{ €})] \simeq 0.093$.

change. Because the abolition of the baby bonus implies an increase in the price of a further child, prospective parents had a financial incentive to move their baby plans forward. From the time of the announcement of the elimination of the birth benefit, the time window during which a baby would have to be conceived in order to still get the birth benefit was very short, however.

In fact, we can be quite precise regarding the length of this time window because the duration of gestation is recorded in the birth statistics (more details are given in section 3.3 below). In the time period considered (i.e. the period from July 1990 until December 2006), the length of a pregnancy shows an approximately normal distribution, with a mean duration of 276 days and a standard deviation of about 14 days. The abolition of the bonus was definitely announced on March 7, 1996. After a mean pregnancy duration of 276 days, birth would take place on December 8 at the earliest. The potential response time for women with average pregnancy duration therefore lasts 23 days (i.e. December 31 - December 8) – a little bit more than three weeks. The corresponding 90% confidence interval ranges from 266 days to 287 days, implying that approximately 90% (10%) of all conceptions from March 19 (April 9) can be expected to be born before January 1, 1997.⁸ These simple calculations make it clear that the window of opportunity was short, and that prospective parents thus had to respond quite immediately if they wanted to still be eligible for the bonus after the abolition had been made public.

3.3 Data and Descriptives

3.3.1 Data Source

Our empirical analysis relies on individual birth records from the Austrian birth statistics (“Geburtenstatistik”), covering all births from 1971 until 2006. In addition to information on year and month of birth, the data also contain some information on parental characteristics (such as age, education, marital status, labor market status, religion, and nationality) and, beginning in 1984, some health measures for the newborn child (such as weight, length, and Apgar score⁹). Moreover, information re-

⁸Note that it is possible that some couples already tried to conceive after February 2, 1996, even though there was confusion about the exact date of abolition until March 7, 1996. It is therefore still possible to find an increase in births before December 1996. See also footnote 12.

⁹The Apgar score is used to assess the health of a newborn immediately after birth. In our data, the Apgar score one, five, and ten minutes after birth is recorded. The Apgar score assesses five different categories (heart rate, breathing, muscle tone, reflex response, and skin color) with a score between zero and two each, where the scores are simply added up. Low values on the score are indicative of poor health. In the regression reported in section 3.5.2 below, we use the average of a child’s score 1, 5, and 10 minutes after birth.

garding the implemented birth procedure is recorded in the data from 1995 onwards.

3.3.2 Sample Period(s)

Our baseline sample period basically covers the period from July 1990 until December 2006. We start the sample period in July 1990 because another major family policy reform took effect on that specific date (the reform basically involved a massive extension of the duration of parental leave benefits).¹⁰ The sample period is considerably shorter, however, when we focus either on newborn's health or on birth procedure due to data availability.

As we will show below, however, the exact length of the sample period does not appear to have any substantial impact on our results. In fact, our estimates of the policy impact turn out to be robust across a wide range of alternative sample periods (as will be shown in section 4.5.2 below).

3.3.3 The Monthly Birth Count, 1990–2006

Figure 3.1 shows both the observed and the de-trended absolute number of monthly births from January 1990 to December 2006.

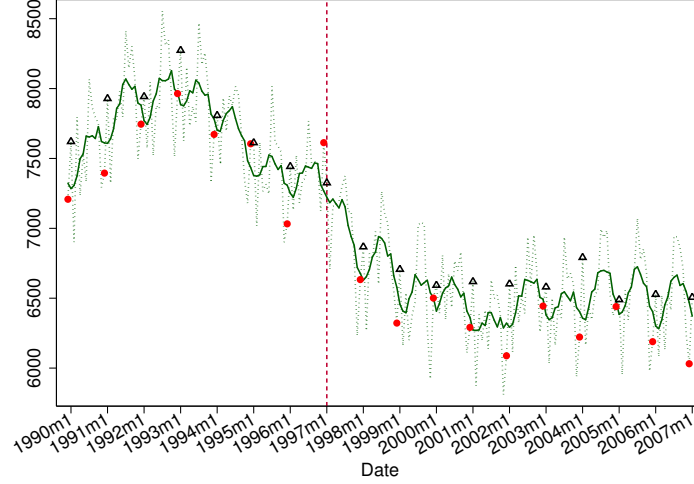
Panel (a) shows the absolute number of monthly births, with dots (triangles) indicating the number of births in December (January) in any given year. Two specific features stand out clearly. First, there is a strong non-linear trend in the number of births, with a pronounced hump shape in the 1990s (presumably reflecting the large immigrant influx from the Balkan countries at that time) and a flattening afterwards. The number of monthly births increased from about 7500 births per month in the early 1990s to a high of somewhat more than 8000 births per month in the mid 1990s. The number of births began to decrease again at the end of the 1990s, when the number of births seems to have stabilized at about 6500 births per month. The second outstanding feature is the existence of a pronounced cyclical pattern within any given year. Within each year, many more children are born in the middle rather than at the end of the year.¹¹ Even more striking is the fact that the number of children

¹⁰It is worth mentioning that our basic sample period covers, besides the austerity package and the abolition of the baby bonus, two other major policy changes in family law that were made public in August 2001. First, parental leave duration was extended from 18 to 30 months for all mothers who were on maternity leave during August 2001, gave birth after July 2000, and earned no more than 14,600€ per year. A second reform was enacted in January 2002 and decoupled eligibility to maternity leave benefits from any prior work requirement, thus extending eligibility to self-employed women and mothers not in the labor force. We control for these policy changes by including appropriately defined indicator variables in the regressions that are based on sample periods covering these policies.

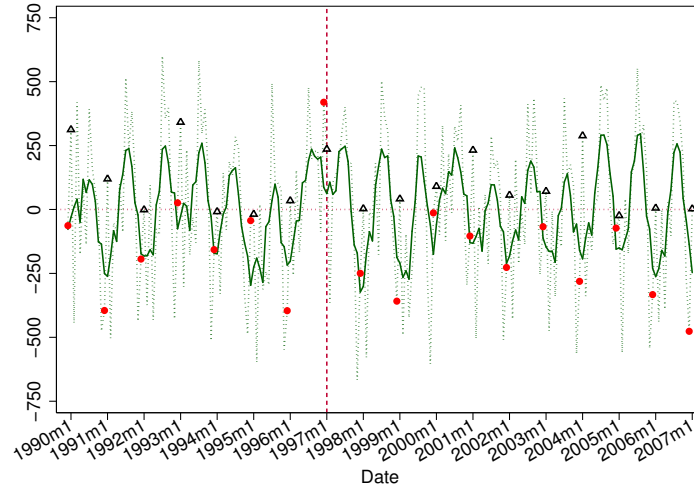
¹¹There are basically two explanations for the seasonal pattern in birth timing. First, there are

born in December over the whole period considered never exceeds the number of births in January – except in December 1996, the month just before the birth benefit was effectively abolished.

Figure 3.1: Actual and De-trended Number of Monthly Births



(a) Observed absolute number of monthly births



(b) De-trended number of monthly births

Notes: Panel (a) shows the observed number of monthly births. The dotted line shows the actual number of monthly births, the solid line displays the three-month moving average. Panel (b) shows the de-trended number of monthly births (de-trending is done using a Hodrick-Prescott filter with default smoothing parameter). The solid line represents the two-month moving average. The dots (triangles) indicate the actual or de-trended number of births in December (January) of each year.

seasonal fluctuations in marriages which may lead to fluctuations in births. In fact, marriage seasonality in Austria matches the seasonal pattern in births if newlywed couples immediately stop using contraceptives with the intent of conceiving. A second explanation are parental preferences regarding the month of birth (Rodgers and Udry, 1988). See also Buckles and Hungerman (2008) for a detailed discussion of both causes and consequences of seasonality in births.

We remove the time trend from the data in panel (b) and thus make the monthly cyclical in births more evident. In each year, the number of births is lowest at the end of the year and highest in the middle. The year 1996 clearly stands out regarding the number of children, especially those born in December. In fact, the de-trended number of births in December 1996 (463) corresponds to the 95th percentile of the distribution of de-trended number of births over the entire period and to the 1st percentile of the distribution of the de-trended number of births in December.

3.4 The Fertility Response

3.4.1 Estimating the Fertility Response

A couple of issues have to be considered when estimating the fertility response following the announcement of the abolition of the baby bonus. First note that we have to rely on the longitudinal patterns in the number of births to estimate the policy impact on the birth count. Because there is no control group available, the outcome in the absence of the policy change needs to be predicted using regularities in the data before and/or after the policy change. Another issue is that there may be a permanent effect from the abolition of the baby bonus on fertility behavior. This implies that we should be cautious when using, or potentially try to do without, data from after the baby bonus has been abolished (i.e. data after December 1996). Second, figure 3.1 suggests that we should try to model a flexible time trend in the number of births when using a longer sample period. However, it turns out that fitting a flexible time trend becomes somewhat difficult when only using data from before the policy change because observations at the boundary of the sample period have a strong impact on the estimated time trend in case of a nonlinear trend (and thus on the estimate of the fertility response as well). We use two distinct empirical strategies to cope with these issues in the following.

Our first empirical strategy only uses data from before the policy change until (and including) October 1996, but refrains from fitting a flexible time trend.¹² At the same time, we want to focus on a time period within an approximately linear time trend

¹²November 1996 births are excluded as well because, according to the distribution of the pregnancy duration, about 5% of responding mothers who conceived immediately after announcement delivered before November 28, 1996. This follows from the 90% confidence interval that ranges from 266 to 287 days of pregnancy. More importantly, it turns out that the initial confusion about the exact date of the abolition was less pronounced than our reading of the newspapers suggested (see section 3.2.2 again) and that many prospective parents must have known the date of the abolition already before the first week of March. Indeed, we already find a substantial, and statistically significant number of additional births in November 1996 (results not shown). This implies that our main estimates based on births in December 1996 unambiguously represent a lower bound on the overall fertility effect.

in the number of births. A simple visual inspection of the observed number of births, as in figure 3.1, suggests that there is a linear time trend in the monthly number of births from about 1992 onwards. We thus regress the absolute (or, alternatively, the log) number of births on a linear time trend and a series of dummies for calendar month, denoted by γ_m , on a sample period of varying length

$$b_t = \alpha + \beta t + \gamma_m + \varepsilon_t, \quad (3.1)$$

with $t \in \{T, T+1, \dots, 1996.9, 1996.10\}$. We let T be equal to either 1992.11, 1993.11, or 1994.11. Thus the sample period consists of either 48, 36, or 24 months (in section 4.5.2, we show that our results are also robust to alternative sample periods).

In the second strategy, we use observations from both before and after the policy change, implicitly assuming that only those parents move birth forward who originally wanted to give birth in the time period that we omit from the sample period. Using observations from after the policy change as well makes it possible to use a much more flexible form for the time trend in the number of births

$$b_t = \alpha + \beta \kappa(t) + \gamma_m + \delta I(t \geq 1997.1) + \varepsilon_t, \quad (3.2)$$

with $t \in \{1990.7, 1990.8, \dots, 1996.9, 1996.10, T, T+1, \dots, 2006.11, 2006.12\}$. That is, t runs from July 1990 to December 2006 in this case, but we leave out a period of varying length in the middle of the sample period, running from 1996.11 to T . We let T equal either 1997.1, 1998.1 or 1999.1, and thus the period that is left out from the analysis correspond to either 2, 14, or 26 months. In this second scenario we allow for a flexible time trend in the number of births, using a fourth-order polynomial in t , denoted by $\kappa(t)$. We also allow for the possibility of any permanent effect of the abolition of the bonus on the number of births by including a dummy variable that takes on the value of 1 if t is equal to or greater than 1997.1. Thus δ will capture any permanent fertility effect of the abolition of the bonus.¹³

For either strategy, we then use the estimates from the above regressions in a second step to make an out-of-sample prediction of the number of babies that would have been born in December 1996 in the absence of the policy change, denoted by $\hat{b}_{1996.12}$. The difference between the observed and the predicted number of births in December 1996,

$$b_{1996.12} - \hat{b}_{1996.12}, \quad (3.3)$$

¹³Similarly, we also include additional dummies for the other major policy changes that were implemented during the sample period (see footnote 10 for details).

is our estimate of the impact of the announcement of the abolition of the baby bonus on the number of children born, relative to the number of children we would have expected in the absence of the policy change, or early announcement, alternatively.

Table 3.1: Fertility Responsiveness

| Panel A: Observations from before the abolition only, linear time trend | | | |
|--|-----------------------------------|-----------------------------------|-----------------------------------|
| Sample period | 1994.11-1996.10 | 1993.11-1996.10 | 1992.11-1996.10 |
| Number of births | | | |
| Residual Dec 1996 | 486.625*** (46.648) | 562.583*** (109.883) | 591.333*** (114.649) |
| Number of births Dec 1996 | 7613 | 7613 | 7613 |
| Number of observations | 24 | 36 | 48 |
| Adjusted R-Squared | 0.739 | 0.718 | 0.838 |
| p-value (F-statistic) | 0.000 | 0.000 | 0.000 |
| Log number of births | | | |
| Residual Dec 1996 | 0.066*** (0.006) | 0.075*** (0.014) | 0.078*** (0.014) |
| Number of log births Dec 1996 | 8.938 | 8.937 | 8.938 |
| Number of observations | 24 | 36 | 48 |
| Adjusted R-Squared | 0.732 | 0.715 | 0.839 |
| p-value (F-statistic) | 0.000 | 0.000 | 0.000 |
| Panel B: Observations from before and after the abolition, nonlinear time trend | | | |
| Sample period | 1990.7-1996.10 &1997.1-2006.12 | 1990.7-1996.10 &1998.1-2006.12 | 1990.7-1996.10 &1999.1-2006.12 |
| Number of births | | | |
| Residual Dec 1996 | 698.290*** (129.264) | 678.489*** (124.48) | 663.846*** (123.07) |
| Number of births Dec 1996 | 7613 | 7613 | 7613 |
| Number of observations | 196 | 184 | 172 |
| Adjusted R-Squared | 0.956 | 0.960 | 0.962 |
| p-value (F-statistic) | 0.000 | 0.000 | 0.000 |
| Log number of births | | | |
| Residual Dec 1996 | 0.097*** (0.018) | 0.094*** (0.018) | 0.091*** (0.017) |
| Number of log births Dec 1996 | 8.938 | 8.937 | 8.938 |
| Number of observations | 196 | 184 | 172 |
| Adjusted R-Squared | 0.952 | 0.957 | 0.960 |
| p-value (F-statistic) | 0.000 | 0.000 | 0.000 |

Notes: *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively. In Panel A, the time trend is assumed to be linear and the underlying sample period varies between two (column 2) and four (column 4) years prior to the policy change. In Panel B, the time trend is assumed to follow a fourth-order polynomial, and a dummy variable for each major policy change is included (1997.1, 2000.7, 2002.1). The underlying sample period covers observations from 1990.7 to 2006.12, but leaves out a period in between which is potentially affected by the policy change.

Table 3.1 shows results for both the absolute number of births and the log number of births and for the two different strategies outlined above.¹⁴ Panel A shows the resulting estimates when using data from before the policy only, but for three different sample periods, combined with a simple linear time trend in each case. Depending on the length of the sample period, estimates of the additional number of births in December 1996 range from 487 to 592 births. In all three cases, the estimate of the extra births turns out to be statistically significant. In relative terms, the estimates imply that about 6.8% ($= 100\% \cdot [487/(7613 - 487)]$) to 8.4% ($= 100\% \cdot [592/(7613 - 592)]$) additional children were born due to the announced abolition of the bonus. We get very similar estimates when using the log number of births as dependent variable, as shown in the lower part of panel A. Relative effects in this case range from about 6.8% ($= 100\% \cdot [\exp(0.066) - 1]$) to 8.1% ($= 100\% \cdot [\exp(0.078) - 1]$).

It turns out that our alternative strategy yields very similar estimates, as shown in panel B. Depending on the sample period, estimates range from a low of 664 (9.5%) to a high of 698 (10.1%) extra births in December 1996. Using data from after the policy change as well thus yields somewhat larger estimates than those we obtain when we only use data from before the policy change, but the point estimates based on the two strategies are in fact not statistically different from each other. Again, using the log number of births yields very similar quantitative implications.

3.4.2 Robustness

We first test the sensitivity of our baseline results with respect to (additional) variations in the sample period. Remember that when only using observations from before the abolition of the bonus, our baseline model uses either 24, 36, or 48 months prior to the policy change in order to predict the December 1996 birth count. Panel A of table 4.8 shows the resulting minimum and maximum estimate of the fertility response when we vary the length of the observation period, in steps of one month, from 24 to 48 months. Estimates turn out to be robust to this variation in the sample period. The resulting minimum (maximum) estimate equals 417 (622) births, an estimate well within the range of our baseline estimates. The same conclusion applies to the range of estimates when using the log number of births as dependent variable.

When using observations from both before and after the policy change, our baseline result basically relies on the whole observation period from July 1990 to December 1996, but excludes a period in between of varying length. In contrast to the baseline

¹⁴We also re-ran our baseline regressions using the total fertility rate as dependent variable. Results turn out to be qualitatively similar.

specification, panel B of table 4.8 holds the omitted period fix (1997.1–1999.1), but varies the length of the sample period before and after, from a minimum of 24 months to a maximum of 76 months (both before and after the omitted period). Again, estimates turn out to be surprisingly robust across the various sample periods. The minimum (maximum) estimate among all estimates is equal to 309 (684) additional births in December 1996. As above, we find a quantitatively similar pattern of estimates when modeling the log number of births instead of the absolute number of births.

Table 3.2: Robustness

| Dependent variable | Number of births | | log(Nr of births) | |
|--|-------------------------|-------------------------|---------------------|---------------------|
| | Min | Max | Min | Max |
| Panel A: <i>Observations from before the abolition only, linear time trend</i> | | | | |
| Residual December 1996 | 416.800*** (65.384) | 622.250*** (114.179) | 0.056*** (0.008) | 0.082*** (0.015) |
| Number of (log) births Dec 1996 | 7613 | 7613 | 8.937 | 8.937 |
| Number of observations | 25 | 40 | 25 | 40 |
| Adjusted R-Squared | 0.667 | 0.814 | 0.661 | 0.812 |
| p-value (F-statistic) | 0.000 | 0.000 | 0.000 | 0.000 |
| Panel B: <i>Observations from both before and after the abolition, nonlinear time trend</i> | | | | |
| Residual December 1996 | 308.955*** (102.14) | 683.588*** (140.048) | 0.041*** (0.015) | 0.095*** (0.019) |
| Number of (log) births Dec 1996 | 7613 | 7613 | 8.937 | 8.937 |
| Number of observations | 61 | 81 | 61 | 65 |
| Adjusted R-Squared | 0.941 | 0.909 | 0.937 | 0.863 |
| p-value (F-statistic) | 0.000 | 0.000 | 0.000 | 0.000 |
| Panel C: <i>Placebo regressions, de-trended number of births</i> | | | | |
| | True policy change | | Placebo regressions | |
| | Y:1996 | Y:1995 | Y:1994 | Y:1993 |
| Residual December Year Y | 455.025*** (143.074) | −182.784 (150.167) | 72.090 (135.771) | 68.241 (145.206) |
| Number of births Dec Y | 7613 | 7232 | 7605 | 7672 |
| Number of observations | 48 | 48 | 48 | 48 |
| Adjusted R-Squared | 0.648 | 0.669 | 0.732 | 0.741 |
| p-value (F-statistic) | 0.000 | 0.000 | 0.000 | 0.000 |

Notes: *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively. Panel A uses a linear time trend with a sample period that varies between two and four years prior to the policy change. The model in panel B assumes a time trend that follows the fourth polynomial. The sample period varies between 24 and 76 months before and after the policy change, while the omitted period in between is held fix (1997.1–1999.1). Panel C shows results from several placebo regressions based on the de-trended series of monthly births. See main text for details.

As an additional robustness check, we apply an alternative two-step procedure. In a first step, we de-trend the whole time series using a conventional Hodrick-Prescott filter. We then regress the de-trended number of births on a series of monthly dummies in the second step.¹⁵ As in the baseline model, we use the four foregoing years to predict the de-trended number of births in December 1996. Moreover, this exercise is not only done for the real policy change but also for hypothetical policy changes in December 1993, 1994 and 1995. Panel C of table 4.8 presents the results. The first column shows that the impact of the abolition is estimated to amount to 455 additional births. This estimate is slightly smaller than the estimate obtained by the baseline model, but it is well within the estimated range of estimates from panel A above (i.e. the estimates are not significantly different from each other).¹⁶ The results presented in the subsequent columns show estimates of the residual number of births in the hypothetical scenario that the policy change happened one, two or three years earlier than it actually did. It is immediately evident that none of the placebo regressions yields a residual that is statistically different from zero, underlining our argument that the announced abolition of the baby bonus increased fertility in the short-run.

3.4.3 Conception (Re-)Scheduling vs. Timing of Births

Thus far we have ignored the fact that we expect to see extra births in December 1996 for two very distinct reasons. First, as we have discussed above, there was a short window of opportunity of about three weeks during which prospective mothers could try to get pregnant in order to give birth before January 1, 1997 and still get the birth benefit. A second reason, however, may be that women with a due date close the date of abolition could have manipulated the exact day of birth by means of a surgical intervention (i.e. cesarean section). We now try to gain some insight into the effective source of the additional births that we observe in December 1996. To distinguish between the two channels, we now focus on the date of conception, which can easily be derived from the available information on the duration of pregnancy and the date of birth. Note that, because the abolition of the birth benefit was announced after the first week of March, the three week response window falls entirely into the month of March. Hence the comparison of the impact on the number of conceptions

¹⁵Specifically, we run the following regression: $\tilde{b}_t = \alpha + \gamma_m + \varepsilon_t$, where \tilde{b}_t denotes the de-trended number of monthly births and γ_m denotes the inclusion of a full set of monthly dummies.

¹⁶It is actually quite intuitive that the estimate based on the de-trended number of births is smaller because the filter fits the time trend using all observations – including the extra births in December 1996. As a consequence, the time trend is biased upward around the date of the true policy change. This in turn results in a downward biased estimate for the fertility response in December 1996.

in March with the impact on the number of births in December 1996 is insightful in terms of whether conception (re-)scheduling or birth timing is the primary cause of the extra births in December 1996. Analogous to the baseline model, we use data from the preceding 48 months to make a simple prediction of the number of babies that would have been conceived in March 1996 in the absence of the policy change.

Table 3.3: Conception (Re-)Scheduling vs. Birth Timing

| Dependent variable | Nr of conceptions | Log nr of conceptions |
|--|-------------------------|-----------------------|
| Residual March 1996 | 631.041*** (135.143) | 0.083*** (0.017) |
| Number of (log) conceptions March 1996 | 7547 | 8.928 |
| Number of observations | 48 | 48 |
| Adjusted R-Squared | 0.783 | 0.784 |
| p-value (F-statistic) | 0.000 | 0.000 |

Notes: *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively. The dependent variable is the (log) number of conceptions in March 1996. The regression specification assumes a linear time trend. The sample period covers all conceptions within the four years preceding March 1996.

The resulting estimates, shown in table 3.3, imply that 631, or about 9.1% ($= 100\% \cdot 631 / (7547 - 631)$), additional children were conceived in March 1996 on top of what would have been expected in the absence of the policy change. Remember that our baseline model yields an estimate of 591 additional births in December 1996 (see table 3.1) – almost the same number as our estimate for the additional number of conceptions in March 1996. The fact that both results are very much in line with each other suggests that conception (re-)scheduling, rather than birth timing by medical intervention, is the underlying cause of the observed fertility response. Section 3.5.1 below provides additional evidence in line with this result, showing that there is no impact on the fraction of cesarean sections conducted in December 1996. Finally, note that it is likely that the number of couples trying to move baby plans forward is likely to be higher than those 616 who finally succeeded.¹⁷

¹⁷A rough approximation of the total number of responding couples is obtained by multiplying the number of extra births with the probability of becoming pregnant within three weeks. Gnath *et al.* (2003) study the likelihood of spontaneous conception in subsequent cycles for a random sample of German women and find that cumulative probabilities of conception at one, three, six and twelve cycle(s) are, respectively, 38%, 68%, 81% and 92%. A linear interpolation between month zero and month one yields a cumulative probability of conception of 29% at week three, which implies that approximately 2038 ($= 591 / 0.29$) couples were induced to bring their baby plans forward. Relative to the December 1996 birth count that would have occurred in the absence of the policy change, the responsive sample thus amounts to as much as 29% ($= 2038 / (7613 - 591)$).

3.5 Taking Risks for the Bonus?

3.5.1 Birth Complications

We now try to understand whether mothers (un)consciously take health risks for themselves and/or their newborn child when trying to obtain the bonus. We start looking at birth complications. In the following we consider instrumental vaginal delivery mechanisms (forceps delivery, vacuum extraction, and breech delivery) as indication of birth complications, as all three delivery methods involve potential health risks for mother and/or child and are thus applied in emergency situations only. While the former two types of assisted deliveries are used in case of maternal exhaustion, fetal distress, or a combination of both, the latter method is used in labor with a baby in head-down position. For simplicity, we will refer to these instrumental vaginal delivery mechanisms as “birth complications” in what follows. Cesarean section is considered separately because it has been performed upon request more recently for deliveries that could otherwise have been natural, even though it is usually performed only when a vaginal delivery would put baby’s or mother’s life or health at risk. Accordingly, we think that a cesarean section must be viewed as an instrument for deliberate birth timing – in line with the results from Gans and Leigh (2009) and Tamm (2009).

To estimate the impact of the abolition of the bonus on the incident of birth complications, we use basically the same regression specification as above but with the percentage share of birth complications as dependent variable. Because information on birth procedure is only reported from 1995 onwards, we use data from both before and after the policy change to estimate the impact on birth procedures. Specifically, the sample period covers the period from January 1995 to October 1996 and from January 1998 to December 2006.¹⁸

Table 3.4 reports the baseline result for the percentage of overall birth complications (column 2), as well as for single birth procedures (columns 3 to 5). Finally, the last column shows the estimated impact on the percentage of cesarean sections. We find that there is a statistically significant and substantial increase in the percentage of overall birth complications of about 0.9 percentage points in the month prior to abolition of the baby bonus. Note that this corresponds to a relative increase in the probability of experiencing some birth complication by almost 14% ($= 100\% \cdot (0.924/6.778)$). In absolute numbers, the figures imply that 70 additional complications were observed in December 1996. The following three columns show results by individual delivery

¹⁸Running the same model, but excluding either two or three years (instead of one year) in between, yields very similar estimates.

method. Estimates show a significant increase for all but one of the instrumental vaginal birth procedures (forceps delivery). The overall increase in the share of labor complications is thus mainly driven by an increase of breech deliveries and vacuum extractions. In terms of non-vaginal instrumental delivery methods, column 6 reveals an insignificant estimate for the percentage share of cesarean sections, suggesting that women did not use this method to deliberately manipulate the date of birth.

Table 3.4: Birth Complications

| | Any birth complication | Vacuum extraction | Forceps delivery | Breech delivery | Cesarean section |
|----------------------------|---------------------------|----------------------|---------------------|---------------------|---------------------|
| Percentage share of births | | | | | |
| Residual Dec 1996 | 0.924*** (0.260) | 0.463** (0.218) | 0.114 (0.093) | 0.347*** (0.088) | 0.061 (0.517) |
| Share Dec 1996 | 6.778 | 4.282 | 1.274 | 1.222 | 13.740 |
| Nr of observations | 130 | 130 | 130 | 130 | 130 |
| Adjusted R2 | 0.321 | 0.799 | 0.914 | 0.896 | 0.982 |
| p-value (F-stat.) | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

Note: *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively. The dependent variable in column 2 is the overall percentage share of birth complications (corresponding to the sum of columns 3 to 5), while columns 3 to 6 show the estimated impact on single delivery methods. The underlying sample period covers observations within the interval from 1995.1 to 1996.10 as well as the interval from 1998.1 to 2006.12. The time trend is assumed to follow a fourth-order polynomial, and a dummy variable for each major policy change within this period is included (1997.1, 2000.7, 2002.1).

One potential explanation of this finding is that responsive mothers are simply a selected group of mothers. If (some of) the characteristics of these mothers are associated with preexisting conditions encouraging birth complications, such as age at birth, differential fertility responsiveness may mechanically affect the incidence of birth complications.¹⁹ To get an idea of how important compositional changes are in explaining the observed increase in the likelihood of experiencing some birth complication, we run an additional decomposition exercise based on individual-level data (see appendix 3.A for details). The decomposition results suggest that only a small fraction, about 12%, of the observed increase in the likelihood of experiencing some birth complication can be related to observable compositional changes resulting from differential responsiveness. It is clear, however, that we cannot rule out that addi-

¹⁹For example, Rayl *et al.* (1996) show that maternal characteristics like primiparity and older maternal age are associated with an increased risk of breech birth. The Austrian data show a very similar picture: the major determinants for both instrumental non-vaginal and instrumental vaginal delivery are primiparity and older age (results not shown).

tional, unobserved maternal characteristics (such as mothers' health status) explain some of the increased risk of experiencing some birth complication as well.

At the same time, the fact that observable maternal characteristics explain only a minor part of the increased risk of birth complications also opens up the possibility of an alternative explanation. Specifically, we may plausibly think of the babies moved forward as mistimed pregnancies, in the sense that these pregnancies occurred earlier than initially planned or desired, and there is evidence that mistimed pregnancy is associated with increased behavioral and psychological risks.²⁰ First, the most important behavioral risks associated with mistimed pregnancies are smoking, drinking, and diet; and such behavior is known to be associated with complications at birth (Cheng *et al.*, 2009; Cnattingius and Lambe, 2002). Other studies have found that mistimed pregnancies are associated with psychological distress. For example, Orr *et al.* (1997) find that women with a mistimed pregnancy are more likely to show depressive symptoms than women with an intended pregnancy. Similarly, Cheng *et al.* (2009) find that women with a mistimed pregnancy are more likely to suffer from postpartum depression. Increased psychological distress during pregnancy in turn appears to be associated with an increased risk of pregnancy complications (Mulder *et al.*, 2002; Norbeck and Tilden, 1983).

3.5.2 Newborns' Health

We next explore the direct impact on newborns' health using several distinct health measures: the incidence of a preterm birth, low birth weight, length at birth and the Apgar score, which is a measure for quickly assessing the health of a newborn (cf. footnote 9). We expect to find differences in the health of children born in December 1996 for the same reasons as for birth complications. If newborns' health is associated with characteristics of the mother, differences in the health of newborn children may simply result from heterogeneous fertility responses. While compositional changes may have positive or negative effects on newborns' health, the additional behavioral and psychological risks potentially triggered by a mistimed pregnancy are expected to unambiguously harm the health of the newborn.²¹

²⁰In the medical and epidemiological literature, a mistimed pregnancy is usually defined as a pregnancy that occurred earlier than desired (e.g. Cheng *et al.*, 2009). Under normal circumstances, antedating a child is a conscious action and should not be considered a mistimed pregnancy. In our case, however, incentives to antedate a child were increased exogenously while other relevant circumstances (e.g. financial situation, health behavior, workload, size of the apartment) remained unchanged. In such a situation, one may argue that mothers are exposed to similar risks as in the case of a truly mistimed pregnancy.

²¹For example, Pulley *et al.* (2002) find that the mistiming of a pregnancy positively correlates with the probability of a preterm delivery (and low birth weight). They conclude that pregnancies

Table 3.5: Newborns' Health

| | Poor health (Apgar < 7) | Small birth length (< 45 cm) | Low birth weight (< 2500 g) | Premature birth (< 37 weeks) |
|----------------------------|----------------------------|------------------------------------|-----------------------------------|------------------------------------|
| Percentage share of births | | | | |
| Residual Dec 1996 | 0.017 (0.132) | -0.021 (0.212) | 0.078 (0.301) | 0.170 (0.267) |
| Share Dec 1996 | 1.352 | 2.522 | 5.713 | 5.254 |
| Nr of observations | 97 | 97 | 97 | 97 |
| Adjusted R ² | 0.235 | 0.523 | 0.632 | 0.741 |
| p-value (F-statistic) | 0.001 | 0.000 | 0.000 | 0.000 |

Note: *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively. The dependent variable is the percentage share of newborn children that are in poor health (indicated by a low Apgar score), of small birth length, of low birth weight or born prematurely. We use the average of a child's Apgar score 1, 5, and 10 minutes after birth. The sample period in each case covers observations within the interval from 1992.11 to 1996.10 as well as the interval from 1998.1 to 2001.12.

Table 3.5 shows the resulting estimates for four different, more or less direct health measures. All estimates are derived applying the same estimation strategy as in the case of the fertility responsiveness (see section 3.4 above). The sample period covers the four years before and after the abolition of the benefit. The first column shows the effect on the percentage of newborn children with a low Apgar score (i.e. a score lower than 7). The resulting point estimate is small and statistically not different from zero. Similarly, we do not find any negative effect on the likelihood of small birth length, of low birth weight or of a premature birth.²² Overall, we thus find no statistical evidence for any (immediate) negative impact on the health of newborn children – despite the fact that we find evidence of increased labor complications, which would suggest that the abolition of the baby bonus put some children at risk. Of course, this finding thus not rule out the existence of any health effect in the medium or the long run.

that are mistimed by more than a few months may have severe health consequences for both mother and child. Similar results for unintended (i.e. both mistimed and unwanted) births are reported by Orr *et al.* (2008).

²²Note that it may make sense to look at the share of premature births conceived in March rather than born in December 1996. This is because babies conceived within the relevant time window of three weeks following the announcement, when born prematurely, would have been born at most 8.5 months later and thus probably already in November. However, this yields an insignificant estimate as well.

3.6 Conclusions

We study the fertility and health effects preceding the abolition of the Austrian baby bonus on January 1, 1997. Even though the bonus was rather small relative to other family benefits available, it was still worth about 4.5% of the median annual labor income in the year of its abolition. Moreover, because the abolition was made public about ten months in advance, prospective parents not only had a financial incentive but also the possibility to react without the need of medical intervention.

We find that about 8% (roughly 600) more babies were born than in the absence of (the public announcement of) the policy change in December 1996, the month before the abolition of the baby bonus. This effect proves to be robust to a variety of robustness checks. Also, considering the fact that the window of opportunity was quite a short period of about three weeks only, the fertility response appears to be quite large. We also find (re-)scheduling of conceptions rather than direct birth timing (through medical intervention) to be the source of the fertility response. Our analysis of birth procedures further reveals a significant and substantial increase in the fraction of mothers experiencing some kind of birth complications by about one percentage point (a relative increase in the likelihood of about 14%). We calculate that only a small fraction of this increase in birth complications can be attributed to changes in observable maternal characteristics. It thus appears plausible that some part of the unexplained increase in birth complications is caused by an underlying increase in behavioral and/or psychological risks triggered by the mistiming of pregnancies (while the other part is best viewed as being caused by unobserved maternal characteristics). However, the increase in birth complications notwithstanding, we do not find any adverse immediate impact on newborns' health.

On a more general level, our results illustrate that announcement effects may be an important issue in health policy reforms, and the abolition of the Austrian birth benefit clearly shows that even relatively small changes in financial incentives may trigger substantial behavioral responses. Policy makers should thus be aware that not only a policy (reform) itself, but also the public announcement of its abolition (or introduction) may have an impact on individual behavior. Second, our results also suggest that policy announcements may lead some individuals to make bad choices in the sense that they unconsciously take health risks in return for a short-run financial benefit. Even though we cannot pin down the importance of this mechanism exactly, it seems fair to say that dealing with announcement effects appears to be especially important in the context of health policy, both in the planning and implementation as well as in the ex-post evaluation of specific policy measures.

3.A Appendix

Decomposing the Increase in Birth Complications

To explore whether the increase in birth complications is due to unobserved stress or due to selection, we perform a simple regression-based decomposition analysis based on individual-level data (Blinder, 1973; Oaxaca, 1973). The goal of this exercise is to determine the impact of selective fertility responses on the likelihood of some birth complication. For the decomposition analysis we simply compare mothers who gave birth in December 1996 with mothers who gave birth in December 1995. Table A.1 shows the results (note that the two columns differ only in the weighting scheme used for the decomposition).

Table A.1: Oaxaca-Blinder Decomposition

| Dependent variable | Any birth complication | |
|---------------------------------------|-----------------------------------|----------------------------------|
| Prediction December 1996 | 0.0736*** (0.0063) | 0.0736*** (0.0063) |
| Prediction December 1995 | 0.0633*** (0.0063) | 0.0633*** (0.0063) |
| Difference | 0.0104 (0.0069) | 0.0104 (0.0069) |
| <i>Decomposition</i> | | |
| Explained | 0.0014** (0.0007) [13.517%] | 0.0012* (0.0007) [11.359%] |
| Unexplained | 0.0090 (0.0088) | 0.0092 (0.0090) |
| Number of observations: December 1996 | 7,613 | 7,613 |
| Number of observations: December 1995 | 7,032 | 7,613 |
| Weights | Dec. 1995 | Dec. 1996 |

Notes: *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively. Robust standard errors are in parentheses. The dependent variable is a binary variable indicating any birth complication.

In line with the corresponding results from table 3.4, the upper part of table A.1 documents a difference in the likelihood of experiencing some birth complication of about one percentage point (in table A.1, however, the difference is not statistically significant). The lower part of the table shows the decomposition results, revealing that about 12-14% of the observed difference in the probability of some birth complication is explained by differences in observed maternal characteristics between the

two groups of mothers. Consequently, 86-88% of the difference remains unexplained.

The extent to which the unexplained part of the increase in birth complications is driven by an underlying increase in behavioral and/or psychological risks from the mistiming of births depends on whether the included variables describe maternal characteristics comprehensively. If the omitted variables are correlated with responsiveness to the incentive and therefore with group affiliation, then the unexplained part of the decomposition might capture not only increased behavioral/psychological risks, but also other unobserved group differences. In particular, if responding mothers are of poorer health than average mothers and poorer health (which is not observed) is positively correlated with the probability for labor complications, then the unexplained part of the differential would be upward biased and therefore not only capture increased behavioral/psychological risks triggered by mistimed pregnancies. A second issue is whether the included characteristics are affected by the behavioral/psychological risks themselves, which seems unlikely in our setup, however).

Acknowledgments

We thank Christian Dustmann, Rafael Lalive, Josef Zweimüller, as well as seminar participants in Brixen, Engelberg and Zurich for helpful comments and suggestions. We also thank Sandro Favre and Philippe Ruh for great research assistance. An earlier version of this chapter has previously been circulated as “Financial Incentives, the Timing of Births, Birth Complications, and Newborns’ Health: Evidence from the Abolition of Austria’s Baby Bonus”. Financial support from the Austrian Science Fund (FWF) is gratefully acknowledged (S 10304-G16: “The Austrian Center for Labor Economics and the Analysis of the Welfare State”).

CHAPTER 4

Cultural Distance and Attitudes Towards Immigration: Evidence from Swiss Voting Results

“Tolerance, inter-cultural dialogue and respect for diversity are more essential than ever in a world where peoples are becoming more and more closely interconnected.”

Kofi Annan, Former Secretary-General of the United Nations.

4.1 Introduction

Globalization, including the internationalization of labor markets, enhanced transport, and communication technology, accompanied by increasing economic differences between countries has opened up the possibilities for immigration from less developed countries. As a consequence, global international migration is increasing not only in scale but also in the cultural diversity of groups involved in that movement (Hugo, 2005; OECD, 2001). As a result, the host countries have to cope with increased levels of social and cultural diversity and thus face new challenges that may not only lead to additional costs but also cause social tensions or even stir up anti-immigration sentiments.

In this chapter, I shed light on the impact of immigration on anti-immigration attitudes from a novel perspective. First, instead of using attitudinal survey data, I take advantage of Switzerland’s system of direct democracy, where citizens vote regularly on immigration issues; I transform these voting results into a share of anti-immigration

votes and use them as a direct measure of attitudes towards immigration. Second, I focus on understanding how the cultural distance between the immigrant and the resident population affects this relationship. For this purpose, I classify immigrants into different cultural groups based on findings from the World Value Surveys (WVSs), and calculate culture-specific immigrant shares. I consider two groups of immigrants for the main part of the analysis. In particular I differentiate between nationalities with a values system similar to the Swiss (culturally similar immigrants) and nationalities without one (culturally different immigrants).

In summary, I investigate whether and to what extent natives' attitudes toward immigration, measured by the proportion of anti-immigration votes, are causally related to the culture-specific immigrant shares of the locality in which they live. The hypothesis is that Swiss citizens residing in municipalities with a higher immigrant share are more likely to vote for a more restrictive immigration policy, and that this effect is more pronounced the larger the fraction of culturally different immigrants. The idea behind this is that immigrants with different cultural values and beliefs may be perceived as threatening the national culture such as language, religion, or way of life in general.

In order to explore this question, I combine outcomes of national votes about immigration between 1970 and 2010 on municipality-level with Swiss census data. To take endogeneity of the place of residence of both natives and immigrants into account, I follow Dustmann and Preston (2001) and instrument the immigrant share in the municipality with the immigrant share in the local labor market. Identification therefore relies on the assumption that an individual's mobility is geographically limited by exogenous factors like family and friends. In other words, the assumption is that racially motivated mobility only exists within and not between local labor markets. Exploiting information on individuals' mobility, I analyze residential relocations by natives in response to immigration, and provide evidence that strongly supports the validity of the identifying assumption. The local labor market immigrant share can thus be regarded as exogenous, while showing a large correlation with the municipal immigrant share. It is therefore likely to be a good instrument, at least in the Swiss context. Moreover, it must be emphasized that, by instrumenting the municipality immigrant share with the one at a larger geographical region, I explicitly allow for residential mobility response of natives across municipality borders.¹

¹The immigration literature typically uses historical settlement patterns of immigrants in a municipality to instrument the current local immigrant share. However, if natives respond to immigrant inflows and move across community borders (as the results from table 4.6 suggest), the IV estimate is likely to be downwardly biased because the negatively affected individuals are more likely to move elsewhere. In contrast, using the local labor market immigrant share as instrument explicitly allows for this type of mobility response.

The central finding of this chapter is that it is not the overall immigrant share per se, but merely the presence of culturally different immigrants that induces Swiss citizens to vote for a more restrictive immigration policy. In fact, the local share of culturally different immigrants turns out to be a significant and sizable determinant of the proportion of anti-immigration votes. Baseline estimates suggest that a one percentage point increase in the share of culturally different immigrants results in a 1.12 percentage point increase in the share of anti-immigration votes. By contrast, the share of culturally similar immigrants does not seem to affect natives' voting behavior at all. This result also holds for a finer cultural classification of immigrants as well as for each single vote. An additional interesting finding is that the local share of culturally different immigrants is a likewise important determinant of the percentage of right-wing voters, implying a tight link between Swiss citizens' voting behavior in plebiscites in comparison to elections.

The present chapter can be put into the context of the existing literature as follows. *First*, it adds to the literature on immigration and its impact on immigration policy. However, I am only aware of one other study that investigates the relationship between the local immigrant share and voting results. In this study, Tolbert and Hero (1996) analyze the results about California's illegal immigration initiative and find that both homogeneous counties with predominantly white population as well as counties with large Latino populations strongly support the initiative.² However, these findings do not address the endogeneity of immigration and are therefore not able to establish a causal relationship. Studies that analyze the impact of immigration on election outcomes are closely related. For example, Halla *et al.* (2012) provide evidence in a recent paper for a causal link between immigration in Austria and electing the right-wing party, FPÖ. They find that a rise in the local immigrant share by one percentage point leads to an increase in the share of right-wing votes by 0.4 percentage points. Moreover, their results suggest that it is mainly the presence of medium-skilled immigrants that causes voters to lean to the far-right.

Second, the chapter contributes to the literature that relates the local immigrant share to a measure for attitudes towards immigrants based on attitudinal survey questions. Several sociological studies provide suggestive evidence; see for instance, Martínez i Coma and Duval Hernández (2009); Meuleman *et al.* (2009); Quillian (1995); Schneider (2008); Semyonov *et al.* (2006). The literature concludes that anti-foreigner attitudes are positively related to immigration rates. However, all these studies treat the immigrant share as being exogenous. The only exception I am aware

²The illegal immigration initiative denies social services, non-emergency health care, and education to illegal immigrants and requires public agencies to report suspected illegal immigrants to state and federal authorities.

of is the paper by Dustmann and Preston (2001). Using data from the British Social Attitudes Survey, they find a significant negative effect of the local minority share (immigrants from Africa, Asia, and the Caribbean) on attitudes towards minorities. Although surveys have the advantage that they contain specific attitudinal information, they have also some disadvantages. First, there is evidence that survey data suffer from a “hypothetical bias”. This type of bias arises because individuals have little incentive to reveal their true preferences, knowing that their answer will have no real consequences (see, for example, Miguet (2008)). By contrast, it is likely that voting results reflect true sentiments toward immigration because the result is binding.³ A further advantage related to voting results is that they allow me to make direct statements about the magnitude of the effects.

Third, this chapter is related to the literature that uses survey data in order to investigate potential causes of anti-immigration attitudes. The majority of these studies focus on the economic channel (labor market concerns, welfare concerns), but only little evidence exists on the importance of the non-economic channel (cultural concerns).⁴ The latter indicate that opposition towards further immigration is related to immigrants’ origins in such a way that opposition is larger with increasing ethnic difference (Card *et al.*, 2005; Dustmann and Preston, 2007).

Fourth, by estimating culture-specific effects, this chapter is also related to the literature on immigrant culture and their economic outcomes. For instance, Fernández and Fogli (2009) show that culture is a significant determinant of economic outcomes like female labor supply and fertility. The results of my study actually go further by showing that immigrants’ culture even affect other people’s attitudes towards them.

The structure of the chapter is as follows. Section 4.2 briefly outlines the cultural concept based on the findings of the WVSs and explains the separation of Swiss immigrants into different cultural groups. Section 4.3 describes the nature of my data and presents some descriptive statistics. Section 4.4 presents the empirical model and explains the identification strategy. Section 4.5 presents and discusses the results, and section 4.6 concludes.

³A second reason relates to the political discussion preceding a vote which may enable individuals to make a more informed decision on the issue up for vote than in the case when answering a survey question.

⁴In fact, the evidence on economic determinants is somewhat ambiguous. Some studies suggest that fears about labor market competition strongly shape individual attitudes towards immigrants (Mayda, 2006; Scheve and Slaughter, 2001). Other studies find either no evidence for the labor market channel (Card *et al.*, 2005; Hainmueller and Hiscox, 2010; O’Connell, 2011), or suggest that labor market concerns are not a significant determinant once unobserved cultural values and beliefs are controlled for (Hainmueller and Hiscox, 2007; Müller and Tai, 2010). Far less contentious is the finding that fears related to welfare concerns and public burden play an important role in the formation of attitudes (Dustmann and Preston, 2004, 2007; Müller and Tai, 2010).

4.2 Cultural Values and Beliefs

Since the focus of this chapter lies in evaluating the rule of cultural distance between the resident and the immigrant population in connection with anti-immigration attitudes, I like to first of all clarify what I mean by the term “cultural values”.

I follow the concept of the World Value Surveys (WVS); more precisely, I rely on the influential findings in Inglehart and Baker (2000) who, using the first three waves of the WVS (1981-1999), isolate two major value dimensions and calculate the corresponding country scores. The so-called global cultural map, as shown in figure A.1 in the appendix, visualizes their results. It shows how scores of countries are located on these two dimensions, which are first, traditional vs. secular-rational values and second, survival vs. self-expression values. Thus, the global cultural map measures cultural and not geographic proximity. The first dimension reflects the contrast between countries in which religion, parent-child-ties, and deference to authority are very important and those in which they are not. The second dimension is linked with the rise of a postindustrial society in which an increasing share of the population grows up taking survival for granted and attaches increasing emphasis on subjective well-being and self-expression.⁵

Based on these two value dimensions, I separate Swiss immigrants into different cultural groups. For the main part of the analysis, I only differentiate between two groups; (1) nationalities with a similar value system to that of the Swiss, and (2) nationalities who contrast with the Swiss value system in at least one dimension. In section 4.5.2, I further decompose the latter group depending on whether their value difference is one- or two-dimensional. Table 4.1 presents a summary of the ethnic composition of the two groups along with the characterizing values.

The first group, containing culturally similar immigrants, comprises countries that are, like Switzerland, located at the upper right quadrant of the global cultural map. This group encompasses all historically protestant and catholic non-ex-Communist countries as well as the English-speaking OECD countries. They exhibit relatively high (low) secular-rational (traditional) values on the one side and relatively high (low) self-expression (survival) values on the other side. While the former reflects a high emphasis on gender equality, tolerance of human diversity and high levels of trust rather than traditional roles and absolute norms, the latter reflects values of a postindustrial welfare state.

⁵These two value dimensions have been shown to explain 70% of the total cross national variation in values and beliefs. Another result of Inglehart and Baker (2000) is that the broad cultural heritage of a society - Protestant, Catholic, Orthodox, English-speaking, Confucian, or ex-Communist - leaves an imprint on values that endure despite modernization. See also Inglehart and Welzel (2005)

Table 4.1: Drivers of cultural differences between the two groups of Swiss immigrants

| Characteristic | Swiss and culturally similar societies | Culturally different societies |
|---------------------------------------|--|--|
| Panel A: Cultural heritage: | | |
| Cultural zones | Historically protestant, catholic non ex-Communist Europe and english speaking OECD countries | Historically Communist, Confucian, South Asian, African and Latin American countries |
| Panel B: Economic development: | | |
| Industrialization driven values | Low (high) traditional (secular-rational) values: - high emphasis on gender equality - increasing tolerance of human diversity - relatively high levels of trust | Low (high) traditional (sec.-rat.) values in ex-Communist and hist. Confucian societies High (low) traditional (sec.-rat.) values in South Asia, Africa, Latin America |
| Postindustrial driven values | Low (high) survival (self-expression) values: - good health (high life expectancy) - increasing intellectual resources - high level of subjective well-being - emphasis on self-expression - political activism | High (low) survival (self-expr.) values in ex-Comm. societies, South Asia and Africa Equal importance of survival and self-expression values in the hist. catholic (Europe, Latin America), Confucian societies |
| GDPpc | Median is > 25'000 | Median is < 15'000 |
| Effective democracy | high levels in late 1990s | low levels in late 1990s |

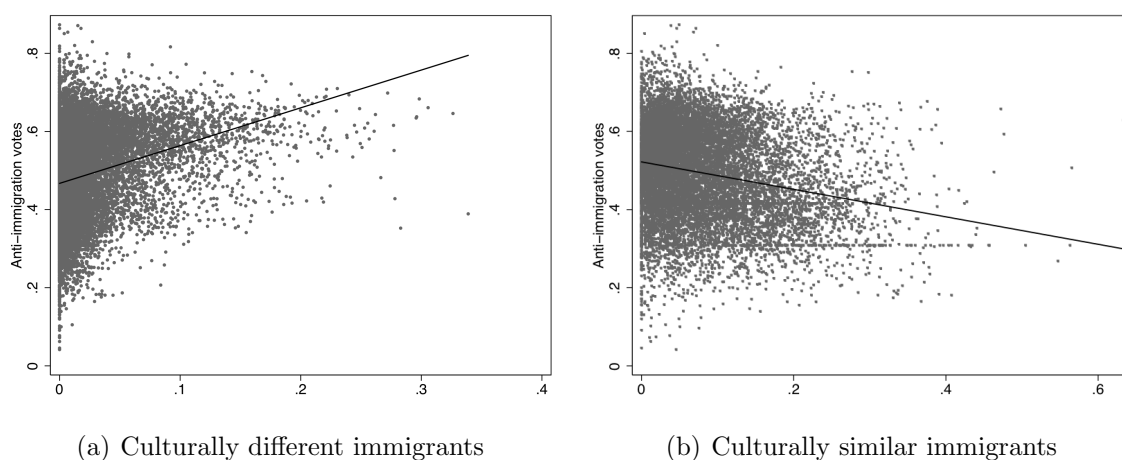
Notes: This table uses results from the World Value Survey (WVS). Especially, it bases on findings by Inglehart and Baker (2000) and Inglehart *et al.* (2003). The per capita GDP is taken from the Penn World Table. It is purchasing power adjusted (real), in USD at 2005 constant prices.

The second group, containing culturally different immigrants, is composed of two slightly different groups. On the one side there are the historically Communist and Confucian countries that do not contrast the Swiss value system in the first, but do so in the second dimension. In particular, the shift away from traditional values towards an emphasis on secular-rational values has already taken place. However, with respect to post-industrial driven values, the historically Communist and Confucian countries differ from Switzerland in that their focus still is on survival rather than on self-expression values. On the other side there are the South Asian, African, and Latin American countries whose values contrast with the Swiss value system in both dimensions, meaning that these societies have experienced neither the post-industrial nor the industrialization driven value shift. As a consequence, they exhibit relatively low levels of trust and tolerance and place low emphasis on gender equality.⁶

⁶The two groups also show different levels of wealth as well as different levels of effective democracy

Figure 4.1 depicts the correlation between the proportion of anti-immigration votes (which I will describe in detail in the following section) and the share of culturally different immigrants in panel (a) and the share of culturally similar immigrants in panel (b). It suggests that cultural distance, as measured by the WVSs, in fact might be a crucial factor in explaining anti-immigration attitudes of the resident population. While there is a positive relationship between the proportion of anti-immigration votes and the share of culturally different immigrants, the reverse picture emerges for culturally similar immigrants.⁷

Figure 4.1: Correlation between immigrant share and share of anti-immigration votes



Note: This figure shows the correlation between the share of culturally different (similar) immigrants and the proportion of anti-immigration votes at municipality-level.

4.3 Data and Descriptives

4.3.1 National Votes on Immigration

The Swiss foundation for research in social sciences (FORS) provides the results of all national votes that took place in Switzerland between 1970 and 1981. Since 1981 official statistics provided by the Federal Statistical Office (FSO) have provided this information. The data are aggregated on municipality-level and contain exact infor-

(Inglehart *et al.*, 2003). While the median 1990 per capita GDP amounts to \$25'000 in advanced industrial countries (Switzerland and culturally similar countries), it does not even amount to \$15'000 in culturally different countries.

⁷This negative correlation may seem strange at first sight, however, it most likely reflects the fact that culturally similar immigrants mainly settle in urban areas where natives also tend to be less hostile to foreigners.

mation on the percentage share of affirmative votes and voter turnout.⁸

I basically use all votes on immigration topics, the only exception being votes with a very clear result, i.e. I exclude votes with a percentage share of affirmative votes of above 70% or below 30%. The underlying reason is that very clear results potentially reflect some kind of general belief rather than a specific attitude toward immigration and, even more importantly, that those who voted against the majority opinion are likely to be a very selected group of individuals. This restriction leads to the exclusion of four votes, leaving me with a sample of 23 votes. Table A.1 in the appendix provides an overview of all the votes.⁹

As apparent from the titles of the single votes, acceptance sometimes means a liberal opinion towards immigration, sometimes it means the opposite. To make votes comparable, I transform the share of affirmative votes into a share of anti-immigration votes. Column six of table A.1 shows whether or not a yes-vote implies support for a more restrictive immigration policy. My dependent variable therefore reflects the share of municipality-level support for anti-immigration policies, which I simply will refer to as anti-immigration votes in the following.

Figure A.2 in the Appendix, showing box-plots of the share of anti-immigration votes for each single votes, illustrates that there is a large variation in the dependent variable between municipalities. In addition, the figure suggests that there is no clear pattern apparent that the common opinion on immigration has changed in an obvious way over time.

4.3.2 Swiss Population Census

The immigrant shares as well as socio-economic and demographic control variables of Swiss citizens are calculated beginning in 1970 from the decennial Swiss population census. The completeness as well as the high data quality allows me to neglect problems of measurement error. In addition to extensive information on individual socio-economic characteristics, the population census reports the country of origin for each foreign born person. It also contains information on individuals' places of residence at the time of the survey as well as five years in the past. This information is

⁸Note that the voting results prior to 1981 for three of total 26 cantons (Swiss states) could no longer be reconstructed on the municipal, but only on the cantonal, level. The cantons concerned are Aargau, Freiburg, and Tessin.

⁹The votes that are excluded are the following: Vote 265 in 1977, which was a popular initiative to restrict immigration. The initiative was rejected with 29% yes-votes. Vote 417 in 1994, and votes 454 and 455 in 1999 were all about alternations in federal law on immigration or asylum policy. All three bills were accepted with more than 70% yes-votes. Nonetheless, these votes are used within the range of a robustness check. See footnote 18 for further information.

insightful in terms of individual mobility behavior and is used to test the plausibility of the identifying assumption in chapter 4.4.2.

Because the census data are only recorded every ten years, I do not have census data for each possible voting year. To still exploit the whole variety of votes, I relate each vote to the year of the nearest census. Column three of table A.1 shows the imputed census year for each vote. However, results turn out to be virtually identical when the data are merged differently.¹⁰

An immigrant is defined as resident without a Swiss passport, but it should be noted that citizenship in Switzerland does not depend on place of birth.

My main regressors are culture-specific immigrant shares aggregated on municipality-level. The calculation is done by first classifying immigrants by country of birth into different cultural groups, and then calculating immigrant shares within these groups. As mentioned above, I confine myself to the following two groups for the main part of the analysis; (1) culturally similar immigrants and (2) culturally different immigrants. Table 4.2 shows how the size and spatial distribution of these two immigrant groups has evolved over time.

Table 4.2: Main explanatory variables

| Decade | Culturally different immigrants | | | | | Culturally similar immigrants | | | | |
|---------|---------------------------------|-------|-------|-------|-------|-------------------------------|-------|-------|-------|-------|
| | Mean | SD | p25 | p50 | p75 | Mean | SD | p25 | p50 | p75 |
| 1970 | 0.014 | 0.010 | 0.006 | 0.012 | 0.021 | 0.158 | 0.082 | 0.108 | 0.154 | 0.202 |
| 1980 | 0.027 | 0.018 | 0.012 | 0.024 | 0.039 | 0.121 | 0.074 | 0.068 | 0.118 | 0.153 |
| 1990 | 0.056 | 0.031 | 0.031 | 0.054 | 0.084 | 0.124 | 0.079 | 0.064 | 0.117 | 0.157 |
| 2000 | 0.089 | 0.051 | 0.046 | 0.088 | 0.134 | 0.115 | 0.071 | 0.061 | 0.105 | 0.149 |
| 2010 | 0.087 | 0.055 | 0.039 | 0.082 | 0.116 | 0.137 | 0.071 | 0.082 | 0.124 | 0.189 |
| overall | 0.057 | 0.049 | 0.018 | 0.040 | 0.088 | 0.131 | 0.077 | 0.072 | 0.123 | 0.173 |

Notes: This table is based on municipality-level data, weighted by the corresponding population size.

There is a strikingly different time trend in the overall size of the two groups. While culturally similar immigrants remained more or less stable in size, the overall size of culturally different immigrants increased sharply over the years. The modest share of 1.4% in the year 1970 increased by a factor of six by the year 2010. Hence, Switzerland experienced a change in the composition of immigrants as it is typical for almost all traditional immigration and OECD countries, namely towards increased

¹⁰I obtain qualitatively identical and quantitatively very similar results when the data are merged differently, namely when votes taking place between 1970 and 1979 are related to the 1970 census, votes taking place between 1980 and 1989 are related to the 1980 census, and so forth.

ethnic diversity. The table also shows that there is considerable variation in the immigrant shares across municipalities.

Immigrant shares calculated at the ms-region-level (ms = “mobilité spatiale”, or spatial mobility) provide the instrumental variable. An ms-region is defined based on the level of economic interactions and commuting patterns of the workforce and thus represents a local labor market. Switzerland is divided into 104 such ms-regions which on average encompass 24 municipalities.¹¹ For the purpose of a robustness check, I also calculate immigrant shares at the canton level. Unlike ms-regions, cantons are subnational entities with far-reaching autonomy in taxation policy or in the educational system, for example. There are 26 cantons, each of which encompasses an average geographical area of about four ms-regions. Table 4.3 reports all spatial units used in my analysis along with their number as well as the mean and median number of municipalities within each unit.

Table 4.3: Spatial units

| Spatial unit | Number of spatial units | Number of municipalities: | | |
|--------------|-------------------------|---------------------------|------|-----------|
| | | Median | Mean | Std. dev. |
| Ms-region | 104 | 20 | 24 | 17 |
| Canton | 26 | 64 | 98 | 102 |
| NUTS2 | 7 | 307 | 363 | 234 |
| Country | 1 | 2544 | 2544 | . |

Note: This table reports for selected spatial units the number as well as the mean and the median of the number of municipalities within each unit.

NUTS2 regions, as listed in row three, represent catchment areas of Switzerland’s seven main economic centers and are included as dummy variables in some of the model specifications. An important feature of NUTS2 regions is that their borders largely represent language borders.¹² Finally, as shown in the last column, the total number of municipalities amounts to 2544. I deal with time-varying municipality structures by creating a balanced panel of municipalities based on a slightly modified version of the territorial boundaries of the year 2000. My final dataset therefore consists of 58’512 observations (23 votes \times 2544 municipalities).

¹¹I use a slightly modified version of the official ms-region classification (with 106 ms-regions) because there are two ms-regions that only encompass one municipality (Davos and Zurich), and which therefore have been combined with its next closest local labor market (Zurich to Limmattal and Davos to Schanfigg).

¹²The definition of the Swiss NUTS2-regions is in compliance with the spatial structures defined by the statistical office of the European Union (Eurostat) and thus allows for direct regional comparisons within Europe. The seven NUTS2-regions are: The lake Geneva region, Espace Mittelland, Northwestern Switzerland, Zurich, Eastern Switzerland, Central Switzerland, Ticino.

4.3.3 Descriptives

Table 4.4 sheds light on the ethnic composition of the two groups of Swiss immigrants. It shows the size of immigrants by country of origin relative to all immigrants for each decade. Looking first at the composition of culturally similar immigrants as shown

Table 4.4: Immigrants by country of origin over time

| | 1970 | 1980 | 1990 | 2000 | 2010 |
|---|-------|-------|-------|-------|-------|
| Panel A: Culturally similar immigrants | | | | | |
| Italy | 0.548 | 0.454 | 0.320 | 0.222 | 0.161 |
| Spain | 0.114 | 0.117 | 0.104 | 0.058 | 0.036 |
| Germany | 0.111 | 0.095 | 0.072 | 0.077 | 0.147 |
| France | 0.052 | 0.052 | 0.044 | 0.043 | 0.054 |
| Austria | 0.042 | 0.035 | 0.025 | 0.021 | 0.021 |
| UK | 0.014 | 0.017 | 0.015 | 0.015 | 0.021 |
| Netherlands | 0.010 | 0.011 | 0.010 | 0.010 | 0.011 |
| USA | 0.011 | 0.010 | 0.009 | 0.010 | 0.011 |
| Greece | 0.009 | 0.010 | 0.007 | 0.004 | 0.004 |
| Belgium | 0.004 | 0.005 | 0.005 | 0.005 | 0.006 |
| Sweden | 0.003 | 0.004 | 0.004 | 0.004 | 0.004 |
| Portugal | 0.003 | 0.021 | 0.092 | 0.098 | 0.119 |
| Canada | 0.002 | 0.002 | 0.002 | 0.003 | 0.004 |
| Liechtenstein | 0.002 | 0.002 | 0.001 | 0.001 | 0.001 |
| Denmark | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 |
| Finland | 0.001 | 0.002 | 0.002 | 0.002 | 0.002 |
| Oceania | 0.000 | 0.000 | 0.000 | 0.000 | 0.003 |
| | 0.929 | 0.837 | 0.715 | 0.577 | 0.606 |
| Panel B: Culturally different immigrants | | | | | |
| <i>Former Communist and Confucian countries</i> | | | | | |
| Former Yugoslavia | 0.023 | 0.066 | 0.144 | 0.250 | 0.176 |
| Turkey | 0.011 | 0.042 | 0.068 | 0.057 | 0.040 |
| Czechoslovakia | 0.012 | 0.016 | 0.005 | 0.004 | 0.007 |
| Rest of Eastern Europe | 0.004 | 0.006 | 0.006 | 0.012 | 0.019 |
| Hungary | 0.011 | 0.007 | 0.004 | 0.003 | 0.004 |
| Poland | 0.001 | 0.003 | 0.005 | 0.003 | 0.007 |
| Romania | 0.000 | 0.002 | 0.003 | 0.002 | 0.003 |
| | 0.063 | 0.140 | 0.236 | 0.331 | 0.256 |
| <i>Africa, Asia, South America</i> | | | | | |
| Asia | 0.002 | 0.007 | 0.023 | 0.043 | 0.065 |
| Africa | 0.003 | 0.008 | 0.016 | 0.029 | 0.044 |
| South America | 0.004 | 0.008 | 0.009 | 0.019 | 0.029 |
| | 0.008 | 0.023 | 0.048 | 0.091 | 0.138 |
| Total | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

Notes: This table shows the relative size of immigrants by country of origin and decade. Panel A shows the composition of culturally similar immigrants, Panel B shows the composition of culturally different immigrants. The latter further distinguishes between countries which differ from the Swiss and culturally similar immigrants in one cultural dimension, and countries which differ in both cultural dimensions.

in panel A, we see that Italians and Spaniards account for the largest share. They mainly entered the country as labor migrants during the 50's and 60's. They accounted for two-thirds of all immigrants in 1970. Following the oil shock in 1973 and the subsequent recession, however, many labor migrants were forced to leave the country. Subsequently, the number of Italians and Spaniards decreased in both absolute and relative numbers. Since the year 2000, immigrants from Germany represent the second largest group. They are followed by immigrants from France, Austria, UK, the Netherlands, and the US who also account for more than one percent of all immigrants and whose shares have remained quite stable over the years. On the contrary, the share of Portuguese has increased sharply over time and has reached a quite remarkable size by now.

Within the group of culturally different immigrants, panel B clearly shows that immigrants from former Yugoslavia make up the largest share. Many of them entered the country on a refugee basis after the start of the Yugoslav wars in the 1990s. Immigrants also increasingly come from Turkey, Asia, Africa and from the rest of Eastern Europe, although to a lesser extent.

Demographic and socio-economic characteristics for Swiss citizens and the two groups of Swiss immigrants are shown in Table 4.5. Looking first at people's place of residence, we see that immigrants are more concentrated in urban areas and in border communities; the latter mainly applies to culturally similar immigrants. Turning to the demographic structure, the most prominent difference between Swiss and foreigners is the age and gender pattern. Immigrants are on average younger and more likely to be male. The former is even more pronounced in the group of culturally different immigrants. The education structure of the two immigrant groups is surprisingly similar. Moreover, the share of individuals with a tertiary education is exactly the same for all groups, even for the Swiss. The education levels, however, may not be fully transformable for more different cultural groups. Further differences appear with regard to individuals' labor market status. Most striking is the fact that culturally different immigrants are five times more likely to be unemployed than both the Swiss and culturally similar immigrants. In contrast, culturally similar immigrants have a ten percentage point higher employment rate than both other groups. Finally, the three groups contrast strongly from each other with respect to their religious affiliation. Among the Swiss, Protestants and Catholics are about equally distributed, while Jews and Moslems do not even make up one percent of the population. Culturally similar immigrants show a very similar pattern, with the only difference being that within this group Catholics markedly outweigh Protestants. The religious composition of culturally different immigrants is completely different. This group only contains

4% Protestants, and furthermore is approximately one-third Catholics and Moslems each, and about one-fourth with another religion or none at all.

Table 4.5: Average characteristics

| | Swiss | | Immigrants, culturally different similar | | | |
|--|-------|--------|---|--------|------|--------|
| <i>Area characteristics (place of residence)</i> | | | | | | |
| Urban area | 0.29 | (0.23) | 0.38 | (0.25) | 0.39 | (0.26) |
| Living in border community | 0.11 | (0.31) | 0.12 | (0.32) | 0.15 | (0.35) |
| <i>Individual characteristics</i> | | | | | | |
| Female | 0.52 | (0.50) | 0.44 | (0.49) | 0.44 | (0.49) |
| Single | 0.43 | (0.49) | 0.45 | (0.49) | 0.44 | (0.49) |
| Married | 0.46 | (0.50) | 0.51 | (0.49) | 0.51 | (0.49) |
| Divorced, widowed | 0.11 | (0.31) | 0.04 | (0.20) | 0.05 | (0.22) |
| Mean age | 38.3 | (22.9) | 27.6 | (16.1) | 32.6 | (18.7) |
| 15 \geq age < 25 | 0.14 | (0.34) | 0.19 | (0.38) | 0.15 | (0.36) |
| 25 \geq age < 35 | 0.14 | (0.34) | 0.25 | (0.43) | 0.22 | (0.41) |
| 35 \geq age < 45 | 0.14 | (0.34) | 0.18 | (0.38) | 0.18 | (0.38) |
| 45 \geq age < 55 | 0.12 | (0.32) | 0.08 | (0.27) | 0.12 | (0.32) |
| 55 \geq age < 65 | 0.11 | (0.31) | 0.03 | (0.16) | 0.07 | (0.25) |
| Age > 65 | 0.15 | (0.35) | 0.02 | (0.12) | 0.05 | (0.21) |
| Mandatory education | 0.33 | (0.46) | 0.50 | (0.50) | 0.50 | (0.49) |
| Secondary education | 0.45 | (0.49) | 0.21 | (0.40) | 0.26 | (0.43) |
| Tertiary education | 0.12 | (0.32) | 0.12 | (0.32) | 0.12 | (0.32) |
| Other education | 0.27 | (0.44) | 0.37 | (0.48) | 0.29 | (0.45) |
| Unemployed | 0.01 | (0.09) | 0.05 | (0.21) | 0.01 | (0.11) |
| Not employed | 0.22 | (0.41) | 0.16 | (0.36) | 0.14 | (0.35) |
| Employed | 0.55 | (0.49) | 0.54 | (0.49) | 0.64 | (0.47) |
| Self employed | 0.13 | (0.33) | 0.05 | (0.20) | 0.05 | (0.22) |
| Family-run shop | 0.04 | (0.19) | 0.02 | (0.13) | 0.01 | (0.08) |
| Trainee | 0.06 | (0.23) | 0.03 | (0.17) | 0.03 | (0.16) |
| Management position | 0.06 | (0.22) | 0.02 | (0.12) | 0.03 | (0.17) |
| Non-management position | 0.69 | (0.46) | 0.79 | (0.40) | 0.85 | (0.35) |
| Rest | 0.03 | (0.17) | 0.09 | (0.28) | 0.03 | (0.17) |
| Not in working-age | 0.34 | (0.47) | 0.26 | (0.43) | 0.24 | (0.42) |
| Evangelic | 0.49 | (0.49) | 0.04 | (0.20) | 0.11 | (0.31) |
| Catholic | 0.43 | (0.49) | 0.35 | (0.47) | 0.78 | (0.41) |
| Jewish | 0.00 | (0.04) | 0.01 | (0.08) | 0.00 | (0.06) |
| Islamic | 0.00 | (0.04) | 0.37 | (0.48) | 0.00 | (0.04) |
| Other | 0.07 | (0.25) | 0.23 | (0.42) | 0.11 | (0.30) |
| Born abroad | 0.07 | (0.24) | 0.83 | (0.37) | 0.76 | (0.42) |

Notes: This table shows average characteristics of Swiss citizens and the two groups of immigrants for the years 1970, 1980, 1990 and 2000.

4.4 Empirical Model and Identification

4.4.1 Econometric Framework

The dependent variable is the share of anti-immigration votes in municipality j and vote v taking place in year t , and thus lies in the interval between zero and one. The model applied in a first step is OLS with the following baseline specification:

$$AIV_{jv[t]} = \alpha + \beta^c I_{jt}^c + \mathbf{Z}'_{jt} \gamma + \mathbf{X}'_{jt} \delta + \psi_{nv[t]} + \varepsilon_{jv[t]} \quad (4.1)$$

The key parameter to identify is β^c , which denotes how the culture-specific immigrant share is related to the proportion of anti-immigration votes. The vector \mathbf{Z}' comprises residential area characteristics such as urbanization rate and whether municipality j borders a foreign country. The vector \mathbf{X}' includes the following aggregated socio-economic characteristics of Swiss citizens; (i) the distribution of labor market status (share of employed, unemployed, not employed); (ii) the distribution of occupational status (share of individuals who are self employed, employed by family member, a trainee, in a management or non-management position); (iii) the distribution of highest educational attainment (share with primary, secondary and tertiary education); (iv) the distribution of marital status (share single, married, and widowed or divorced); (v) average age and age squared; (vi) the gender distribution; (vii) the distribution of religious affiliation (share Protestants, Catholics, Jews, Moslems and people with another religion or none at all) and finally, (viii) the share of foreign born Swiss. The model further includes NUTS2-specific vote fixed effects, denoted by $\psi_{nv[t]}$. In doing so, I eliminate vote(time)-varying unobserved heterogeneity between the seven catchment areas of Switzerland's main economic centers. In other words, I exploit cross-sectional variation across municipalities within the same NUTS2-region. Finally, I weight all observations with the population size of the municipalities. While this will act as benchmark specification, I will additionally provide estimates of alternative model specifications to test the robustness of the results.

Since individuals are free to choose their neighborhoods, the immigrant share in an individuals' municipality is unlikely to be exogenous. It is reasonable to assume that xenophobic natives are unlikely to locate in municipalities with a large immigrant population. At the same time, it seems unlikely that immigrants will settle in areas where they expect racial discrimination. The OLS estimate of the parameter β^c in equation (4.1) is inconsistent and downward biased under these circumstances.

In order to obtain a consistent estimate of the impact of immigration on voting results, I apply an IV approach in a second step. Following Dustmann and Preston

(2001), I instrument the municipal immigrant shares with the immigrant shares in a larger geographical area. More precisely, I use the immigrant shares in the local labor markets (ms-regions) because they are based on commuting patterns and thus best comply with the requirements on a valid instrument, as will become clear in the next paragraph. The first stage is thus given by the following equation;

$$I_{jt}^c = \pi_0 + \pi_1^c I_{lt}^c + \mathbf{Z}_{jt}'\pi_2 + \mathbf{X}_{jt}'\pi_4 + \psi_{nt} + \varepsilon_{jt}, \quad (4.2)$$

with I_{lt}^c representing the culture-specific immigrant share in time t and ms-region l to which municipality j belongs to.

I argue that individuals, when choosing the place of residence, first decide upon a larger geographic region based on family, friends, and employment history. The crucial point is that the choice does not depend on attitudes towards immigration at this stage. Only in the second step, when it comes to decide upon the local neighborhood within the larger settlement region, might the immigrant composition become part of the decision.¹³

Thus, the key identifying assumption is that the immigrant share in the larger geographic region (ms-region) is exogenously determined by factors like family and friends and not influenced by anti-immigration attitudes. Put differently, I assume that racially motivated mobility only exists within but not between local labor markets (ms-regions). I test the validity of this assumption in section 4.4.2 by estimating mobility effects of changes in the municipal immigrant share. The results indeed provide strong evidence in favor of the identifying assumption. This means that the immigrant share in the ms-region is uncorrelated with unobserved factors determining voting behavior. Consequently, I_{lt} is a valid instrument for I_{jt} because sorting within an ms-region, which is that part of the settlement decision that can depend on attitudes, does not alter the overall ethnic composition of that ms-region.

Identification further requires that the instrument has no direct effect on voting behavior. This implies that an individual's attitude toward immigration is assumed to depend solely on the immigrant concentration in the immediate vicinity (municipality) and not on the average immigrant share of the ms-region. I perform different checks to support the plausibility of this assumption. For instance, I use the ms-region immigrant share ten years earlier as alternative instrument because a direct effect is

¹³Notice that high foreign labor market competition could lead to a situation in which even job location and therefore the choice of the local labor market may be endogenous. However, findings from the 2007 Swiss Household Panel show that only 0.85% of the Swiss population changed residence within the last two years for employment related reasons. If we further consider that only a fraction of these 0.85% might have changed their job because of increased foreign competition, this is probably a minor issue in the particular case of Switzerland.

less likely to exist for immigrant shares that existed one decade earlier.

I also apply the instrument that is typically used in the literature on immigration, namely the historical settlement pattern of immigrants in a municipality. This instrument takes advantage of the fact that immigrants are more likely to settle in areas with already existing shares of people with a similar ethnic background. The identifying assumption is that past migration patterns have no effect on current attitudes, which seems quite plausible. However, there are some disadvantages associated with this instrument. First and most importantly, this instrument does not allow for any residential mobility response of natives across municipality-borders. As a result, using historical migration patterns as an instrument is likely to produce downward biased estimates because the negatively affected individuals are more likely to move elsewhere. A further drawback of this instrument is that one loses one time period and, considering that my dataset only covers five decades, this is quite a large loss. Nonetheless, the results might provide useful insights regarding my identification strategy and can be compared with the results of other studies.

4.4.2 Validity of the Identifying Assumption

Identification requires that racially motivated mobility by the voting population, i.e. Swiss citizens, only exist within but not across local labor markets (ms-regions). Since the population census reports individuals' places of residence at the time of the survey as well as five years earlier, I can test the validity of the assumption by estimating the following equation;

$$1(r_t \neq r_{t-5})_i = \alpha + \beta \Delta I_{j0t} + \mathbf{Z}'_{j0t} \gamma_0 + \mathbf{Z}'_{jdt} \gamma_d + \mathbf{X}'_{it} \delta + \psi_{nt} + \varepsilon_{it}, \quad r \in \{j, l\} \quad (4.3)$$

where the dependent variable is an indicator function taking on the value one if individual i moved within the last five years (between t and $t - 5$) (a) across a municipality border and (b) across a ms-region border. The vectors \mathbf{Z}'_{j0t} and \mathbf{Z}'_{jdt} contain residential area characteristics (unemployment rate, urbanization rate, and border municipality) of the origin and destination municipality, respectively. The vector \mathbf{X}' includes individual socio-economic characteristics as specified in equation (4.1). The parameter of interest is β . It denotes how the change in the immigrant share during the last decade in the original resident municipality affects the probability of having left (a) the municipality and (b) the ms-region. The results are shown in Table 4.6.

Table 4.6: Residential Mobility Effects

| | OLS | | Probit | |
|--|-------------------------------|-------------------|-------------------------------|---------------------|
| | Movements across: Municip. | MS-region | Movements across: Municip. | MS-region |
| ΔI_{jot} | -0.056 (0.179) | 0.012 (0.238) | -0.191 (0.562) | 0.072 (1.275) |
| mfxf | | | [-0.067] (0.197) | [0.012] (0.213) |
| $\Delta I_{jot}^{\text{cult. diff.}}$ | 0.367** (0.184) | 0.088 (0.155) | 1.160** (0.571) | 0.470 (0.890) |
| mfxf | | | [0.406**] (0.200) | [0.079] (0.148) |
| $\Delta I_{jot}^{\text{cult. sim.}}$ | -0.273 (0.235) | -0.027 (0.315) | -0.903 (0.737) | -0.130 (1.650) |
| mfxf | | | [-0.316] (0.258) | [-0.022] (0.276) |
| Mean | 0.326 | 0.119 | 0.326 | 0.119 |
| Standard deviation | 0.469 | 0.324 | 0.469 | 0.324 |
| Number of observations (in '000) | 13'047 | 13'047 | 13'047 | 13'047 |
| Adjusted R-Squared | 0.137 | 0.075 | | |
| p-value (F-statistic) | 0.000 | 0.000 | 0.000 | 0.000 |
| Area char. (\mathbf{Z}'_{jo} and \mathbf{Z}'_{jd}) | Yes | Yes | Yes | Yes |
| Individual socio-demogr. char. | Yes | Yes | Yes | Yes |
| NUTS2-specific time FEs | Yes | Yes | Yes | Yes |

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. The dependent variable is an indicator function taking on the value one if the individual moved across municipality or ms-region borders within the past five years and zero otherwise. The main explanatory variable is the change in the share of culturally different and similar immigrants within the last ten years, respectively. Robust standard errors are given in parentheses and are clustered on municipality-level (2544 clusters). Marginal effects are reported in square brackets.

Results suggest that an increase in the share of culturally different immigrants causes natives to leave their home municipality, but they provide no evidence for racially motivated mobility effects across ms-regions. This finding is the same regardless of whether an OLS or a probit model is estimated. I interpret this result as strong evidence in favor of the validity of my identification strategy.¹⁴

It should further be noted that this result also implies that instruments that do not allow for residential mobility response across municipalities, like the historical settlement pattern of immigrants, underestimate the true effect.

¹⁴The analysis for different time periods shows that these effects are slightly decreasing over time (results not shown but can be provided upon request).

4.5 Results

4.5.1 Main Results

Table 4.7 presents the main results. It shows how the overall immigrant share in panel A and the culture-specific immigrant shares in panel B affect(s) the share of anti-immigration votes of an average vote about immigration. Columns two to four use a simple OLS estimator, columns five to seven report results from using the immigrant share(s) at the local labor market - the overall share in panel A and the culture-specific shares in panel B - as instrument(s) for the municipality-level immigrant share(s). All models include vote fixed effects, and estimates are weighted by the population size of the municipalities. I report simple correlations between the immigrant shares and anti-immigration votes in columns two and five (M1), add controls for local area and socio-economic characteristics in columns three and six (M2), and also allow for unobserved heterogeneity within votes between different NUTS2-regions in columns four and seven (M3). The table only reports the estimates of the parameters of interest, the complete estimation results are reported in table A.2 in the appendix.

As can be seen from panel A, results suggest no direct link between the size of the overall immigrant population and the share of anti-immigration votes. There is an initial negative correlation that vanishes, however, the more stringently the model is specified. This is true for both OLS estimates as well as for IV estimates.

Panel B, showing results for the two cultural groups of immigrants, reveals an interesting picture. The relative size of culturally different immigrants turns out to be a significant determinant of the percentage of anti-immigration votes. Both OLS and IV estimates are positive and statistically significant, although the OLS estimates turn out to be much smaller than IV estimates. This is consistent with the idea I pointed out earlier, that the settlement decision of both natives and immigrants depends on attitudes towards each other. Ignoring this type of endogeneity (as it is the case in the OLS approach) results in a downward biased estimate. The reported coefficients can be interpreted as the percentage point change in anti-immigration votes as response to a one percentage point increase in the immigrant share in the municipality. Thus, baseline estimates imply that people tend to give 1.12 percentage point more votes in favor of a more restrictive immigration policy in municipalities which experience a one percentage point increase in the share of culturally different immigrants. Comparing the different model specifications, we see that the effect approximately doubled in size, from 0.49 pp to 1.1 pp, when local area and socio-economic characteristics are controlled for in addition to vote fixed effects. Additional inclusion of NUTS2-specific vote fixed effects leads to a further although slight increase, from 1.1 to 1.2 percentage

points.

Table 4.7: Main Results

| | Dependent variable: Share of anti-immigration votes | | | | | |
|---|---|----------------------|----------------------|----------------------|----------------------|---------------------|
| | OLS | | | IV (2SLS) | | |
| | M1 | M2 | M3 Baseline | M1 | M2 | M3 Baseline |
| Panel A | | | | | | |
| I_{jt} | -0.266*** (0.040) | -0.039 (0.025) | 0.018 (0.027) | -0.437*** (0.058) | -0.157 (0.110) | 0.108 (0.139) |
| F-stat. of excl.instr.(I_{jt}) | | | | 278 | 151 | 94 |
| Panel B | | | | | | |
| $I_{jt}^{cult.diff.}$ | 0.188*** (0.060) | 0.190** (0.084) | 0.140** (0.066) | 0.489* (0.262) | 1.098*** (0.170) | 1.124*** (0.217) |
| $I_{jt}^{cult.sim.}$ | -0.520*** (0.047) | -0.278*** (0.032) | -0.119*** (0.034) | -0.691*** (0.060) | -0.406*** (0.095) | -0.201 (0.136) |
| F-stat.of excl.instr.($I_{jt}^{c.diff.}$) | | | | 107 | 81 | 48 |
| F-stat.of excl.instr.($I_{jt}^{c.sim.}$) | | | | 362 | 235 | 140 |
| Mean | 0.481 | 0.481 | 0.481 | 0.481 | 0.481 | 0.481 |
| Standard deviation | 0.152 | 0.152 | 0.152 | 0.152 | 0.152 | 0.152 |
| Number of observations | 58,512 | 58,512 | 58,512 | 58,512 | 58,512 | 58,512 |
| Area characteristics | No | Yes | Yes | No | Yes | Yes |
| Socio-demografic char. | No | Yes | Yes | No | Yes | Yes |
| NUTS2-specific vote FEs | No | No | Yes | No | No | Yes |

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. The dependent variable is the share of anti-immigration votes on municipality-level. The sample consists of 2544 municipalities and 23 votes (58'512 observations). Robust standard errors are given in parentheses and are clustered on municipality-level (2544 clusters). All models include vote fixed effects, and estimates are weighted by the population size of municipality j at time t .

The second row of panel B, showing estimates for the share of culturally similar immigrants, reveals a quite different picture. First of all, all OLS estimates are significantly negative, probably reflecting positive self selection in the sense that culturally similar immigrants are more likely to settle in urban areas where natives often tend to have more liberal views towards immigration. Applying the instrumental variable approach, which deals with this type of selection, indeed makes this significant negative effect disappear, at least in the preferred model specification. Hence, results suggest no difference in voting behavior between municipalities with low and high shares of culturally similar immigrants.

With regard to my identification strategy, notice that the F-statistic on the excluded instrument(s) is always well above the critical value of 10 as suggested by Staiger and Stock (1997), alleviating weak-instrument concerns. I also would like to draw attention to table A.3 in the appendix which reports coefficients from the first stage and reduced form regressions. The first stage coefficients are all positive and highly significant, validating the positive relationship between the lower (municipality) and higher (local labor market) level spatial unit immigrant share. Finally, the reduced form coefficients turn out to be highly significant for the culture-specific immigrant shares, implying that there is a causal relationship between the proportion of anti-immigration votes and the culture-specific shares of immigrants within a municipality.

Hence, the central finding is that the presence of immigrants that differ from the Swiss with respect to their cultural values and beliefs appears to be a significant and sizable determinant of the proportion of anti-immigration votes, while the presence of culturally similar immigrants does not seem to affect natives' voting behavior at all. Thus it is not the overall immigrant concentration per se, it is solely the presence of culturally distant immigrants that affects Swiss citizens' attitudes. One possible explanation for this finding may relate to natives' fears that immigrants may undermine their national culture, language, religion, or way of life (Card *et al.*, 2005). However, it cannot be ruled out that, aside from cultural concerns (non-economic channel), labor market and welfare concerns (economic channel) might also explain part of the effect. This is because culture is not the only difference between the two groups of immigrants. As can be seen in table 4.5, they also differ with respect to labor market prospects and, even if only to a small extent, in educational attainment.

4.5.2 Robustness

I test the robustness of the baseline results in various dimensions. First, I test the sensitivity with respect to different covariate and fixed effects specifications as well as to the use of alternative instruments. Second, I check whether the above findings also hold for each single vote. A final robustness check relates to the cultural classification of immigrants.

Specification and identification checks

Table 4.8 presents the results for differing covariate and fixed effects specifications in columns M4 to M6 and for alternative instruments in columns M7 to M9. For comparison purposes, the second column simply replicates the baseline results. The

specifications of model 4 to 6 are as follows: Model 4 differs from the baseline model in allowing for interaction effects between gender, age, and education. To be precise, this model includes the gender-age-education distribution with four education categories and seven ten-year age categories. The point estimates appear slightly larger, but in statistical terms they remain unchanged. Model 5 differs from the baseline model in the way it deals with the aggregated nature of the data. While the baseline model weights all observations by the population size of the corresponding municipality, model 5 instead includes the population size and the population size squared (in 1'000) as additional control variables. This leads to a significantly larger effect of the share of culturally different immigrants. Finally, model 6 includes canton-specific instead of NUTS2-specific vote fixed effects. This means that only the variation within a canton is exploited in this case. The resulting point estimate is only half as large, but a look at the F-value of the excluded instrument gives rise to weak-instrument concerns. Therefore, the point estimates may be biased and have to be taken with caution. Overall, the estimates of immigration on voting results turn out quite robust across specifications.¹⁵

Consider now the last three columns of table 4.8. These models use the baseline covariate and fixed effects specifications but apply different instruments. Model 7 uses the 1970 settlement pattern of Swiss immigrants, the disadvantage being that observations of this decade are lost and, even more importantly, that it does not allow for any residential mobility response of natives across municipality borders. The point estimate is 0.7 and thus, in line with expectations, quite a bit smaller compared to the baseline result.

Model 8 again applies the ms-region level instrument, but instead of the contemporaneous immigrant share, it uses the immigrant share of the last decade. The intention behind this instrument is to assess the validity of the exclusion restriction. Remember, I assume that the municipality immigrant share encompasses the full range of influence, meaning that there is no direct effect of the average local labor market immigrant share on voting behavior. Using the past immigrant share as instrument sheds light on this issue, as a direct effect is less likely for immigrant shares ten years earlier. As evident from column M8, the point estimate for the share of culturally different immigrants is 0.92 and thus very similar to the baseline estimate. In statistical terms they are even identical, which in my view is a strong indication in favor of the validity of my instrument.¹⁶

¹⁵As evident from panel B of table A.3 in the appendix, all first stage coefficients show the expected sign. They are also statistically significant, just like the reduced form coefficients.

¹⁶Admittedly, a potential direct effect cannot be completely ruled out because of low individual mobility and relatively high correlation between ms-region immigrant shares over time (0.63 for

Table 4.8: Robustness of the baseline IV results

| Dependent variable: Share of anti-immigration votes | | | | | | | | | |
|---|--|---------------------|---------------------|---------------------|--------------------------|---------------------|---------------------|----|------------|
| | Checks with respect to covariates and FEs specification | | | | Alternative instruments: | | | | |
| | Baseline | M4 | M5 | M6 | $I_{j,1970}^c$ | M7 | M8 | M9 | I_{mt}^c |
| $I_{jt}^{cult.diff.}$ | 1.124*** (0.217) | 1.369*** (0.171) | 2.460*** (0.225) | 0.545*** (0.255) | 0.793*** (0.191) | 0.917*** (0.260) | 2.488*** (0.503) | | |
| $I_{jt}^{cult.sim.}$ | -0.201 (0.136) | -0.184 (0.134) | -0.100 (0.103) | 0.010 (0.201) | -0.381** (0.160) | -0.138 (0.151) | 0.323 (0.304) | | |
| Mean | 0.481 | 0.481 | 0.494 | 0.481 | 0.488 | 0.488 | 0.481 | | |
| Standard deviation | 0.152 | 0.152 | 0.169 | 0.152 | 0.153 | 0.153 | 0.152 | | |
| Number of observations | 58,512 | 58,512 | 58,512 | 58,512 | 53,424 | 53,424 | 58,512 | | |
| Adjusted R-Squared | 0.762 | 0.760 | 0.547 | 0.858 | 0.803 | 0.796 | 0.464 | | |
| F-stat. of excl. instr. ($I_{jt}^{c.diff}$) | 48 | 68 | 130 | 9 | 56 | 51 | 40 | | |
| F-stat. of excl. instr. ($I_{jt}^{c.sim}$) | 140 | 146 | 223 | 52 | 185 | 127 | 91 | | |
| Area and socio-economic characteristics | Yes | Yes | Yes | Yes | Yes | Yes | Yes | | |
| Alternative spec. of socio-economic characteristics | No | Yes | No | No | No | No | No | | |
| NUTS2-specific vote fixed effects | Yes | Yes | Yes | No | Yes | Yes | Yes | | |
| Canton-specific vote fixed effects | No | No | No | Yes | No | No | No | | |
| Not weighted by, but controlled for population size | No | No | Yes | No | No | No | No | | |

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. The dependent variable is the share of anti-immigration votes on municipality-level. Robust standard errors are given in parentheses and are clustered on municipality-level (2544 clusters). All estimates are weighted by the population size at municipality-level with exception of model 5 which instead controls for the population size and the population size squared (in 1000). The first column just replicates the baseline IV results. Models 4 to 6 use different covariate and fixed effects specifications and models 7 to 9 use the baseline model specification but alternative instruments. $I_{j,1970}^c$ denotes historical settlement patterns of immigrants with the same cultural background, I_{it-1}^c corresponds to the culture-specific immigrant share in the ms-region one decade earlier, and I_{mt}^c is the canton-level culture-specific immigrant share.

Finally, model 9 addresses the validity of the identifying assumption. Remember that although I do not find a significant residential mobility effect across ms-regions, the point estimate of the change in the immigrant share is positive, meaning that there still might be a few individuals whose mobility is not entirely restricted to one ms-region, with the consequence that even my IV estimates might be downward biased. Using the canton-level immigrant share as instrument sheds light on this issue because, compared to an ms-region, a canton encompasses a considerably larger geographic region. As shown in column M9, the effect of the share of culturally different immigrants indeed turns out to be significantly larger, indicating that some racially motivated mobility across local labor markets still exists. My baseline IV results can therefore be regarded as lower bound of the true effects of immigration on anti-immigration votes.¹⁷

Functional form

I test for potential non-linearities by re-running the main model while allowing for a flexible functional form. More specifically, I include the culture-specific immigrant shares from the first up to the fourth polynomial. The results provide no evidence for any economically relevant non-linearities in the effect on anti-immigration votes, neither for culturally different nor for culturally similar immigrants. I therefore conclude that the linear model captures the effects well enough.

Single-vote results

All effects discussed so far relate to results of an average vote about immigration. In consideration of the widely varying contents it appears necessary to examine whether the results also hold for each single vote. This is done by re-running the baseline regressions for each vote. Results are depicted in figure 4.2, with the vertical axis indicating the point estimate (with its 95% confidence interval), and the horizontal axis indicating the corresponding vote number. For comparison purposes, the estimate at the far right additionally depicts the overall effect (corresponding to the baseline IV estimate in table 4.7).

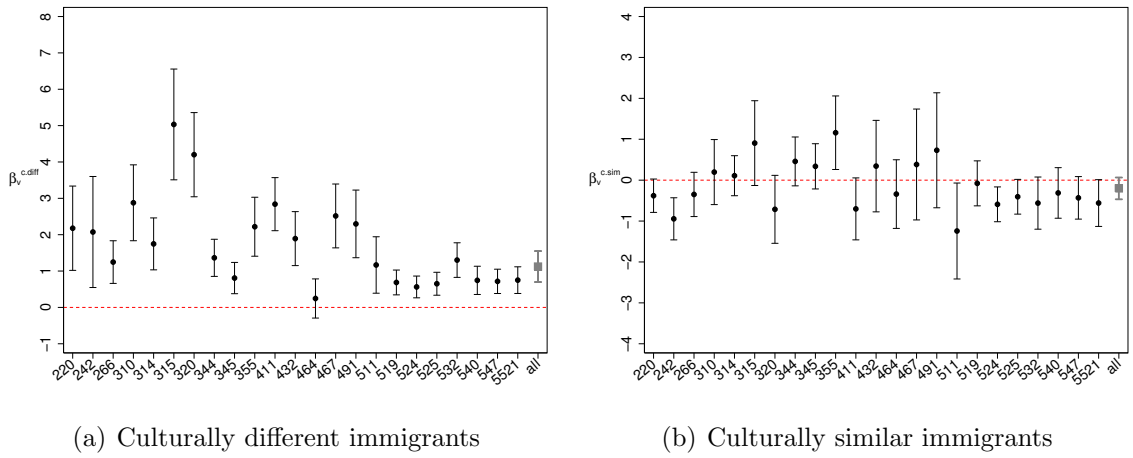
As evident from panel (a), estimates turn out positive for all votes and, with one exception, they are significantly different from zero. In other words, the share of

culturally different immigrants).

¹⁷Even estimates carried out using the canton-level instrument may be downward biased if racially motivated mobility across canton borders also exists. Further notice that first stage and reduced form coefficients are also reported in Table A.3 in the appendix. Again, all first stage coefficients show the expected sign, and both the first stage and reduced form coefficients are highly significant.

culturally different immigrants positively affects the proportion of anti-immigration in all but one vote. The vote where I find no significant effect is vote 464 with the title “federal resolution on bilateral agreements between Switzerland and the EU”. A closer look at the precise content of this particular vote reveals that the vote in fact involved seven bilateral agreements, with the main emphasis being on liberalization and market opening agreements. The free movement of persons was only one of these seven agreements, and more importantly, it was only applied to the EU-8 member states. Under these circumstances, the null effect is no longer surprising.

Figure 4.2: Single-vote results



Notes: This figure shows point estimates (along with the corresponding 95% CIs) from IV regressions of the proportion of anti-immigration votes on the culture-specific share of immigrants. All models use the baseline covariate and fixed effects specification.

There is no clear pattern with respect to the size of the effects. Except for the fact that impacts on votes with the aim of limiting overall immigration are always larger than average, while votes concerning the free movement agreement with the EU seem to have a below-average effect. Apart from that, effects are sometimes above, sometimes below average, independent of whether the vote is on asylum or refugee questions or on naturalization practices.¹⁸

Turning to panel (b), results confirm the previous finding that there is no evidence of a relationship between voting results and the share of culturally similar immigrants.¹⁹

¹⁸I obtain similar results for the four votes that are excluded from the main sample because of a very clear result. More specifically, all coefficients of the share of culturally different immigrants are positive and statistically significant. They vary in size between 0.7 and 2 percentage points.

¹⁹The first stage coefficients again show the expected sign, and the smallest F-value on the excluded instrument lies with a value of 19 sufficiently above the critical value. The reduced form also exists, at least in panel (a).

Overall, this analysis shows that the general picture is quite robust and that the pattern of the pooled cross-section result also holds for each single vote.

Sensitivity to cultural classification

A final robustness check relates to the cultural classification of Swiss immigrants. A look at the global cultural map reveals that in contrast to culturally similar immigrants, there is quite a lot of heterogeneity in cultural values between societies within the group of culturally different immigrants. In fact, the latter includes both societies that differ in one value dimension (i.e. they place high emphasis on survival instead of self-expression values) and societies that differ in both dimensions (i.e. they further attach great importance to traditional rather than secular-rational values).

This section aims to investigate whether the effect of the share of culturally different immigrants is driven by one of these two value dimensions. For this purpose I re-run the baseline model with three groups of immigrants which are (1) culturally similar immigrants (unchanged), (2a) culturally different immigrants in one dimension, and (2b) culturally different immigrants in both dimensions. The ethnic composition of the two subgroups is shown in panel B of table 4.4. It reveals that the group that differs in one value dimension is clearly dominated by former Yugoslavs, whereas the group that differs in both dimensions is quite balanced across South Asia, Africa, and Latin America. Table 4.9 presents the baseline IV results.

Table 4.9: Sensitivity to cultural classification

| | Two groups of immigrants | Three groups |
|---|--------------------------|---------------------|
| $I_{jt}^{\text{cult.diff.}}$ | 1.124*** (0.217) | |
| $I_{jt}^{\text{cult.sim.}}$ | -0.201 (0.136) | -0.200 (0.125) |
| $I_{jt}^{\text{cult.diff.dim1}}$ | | 1.120*** (0.192) |
| $I_{jt}^{\text{cult.diff.dim2}}$ | | 1.127** (0.567) |
| F-value on excluded instrument ($I_{jt}^{\text{cult.sim.}}$) | 140 | 102 |
| F-value on excluded instrument ($I_{jt}^{\text{cult.diff.}}$) | 48 | 108; 11 |
| Number of observations | 58,512 | 58,512 |
| Adjusted R-Squared | 0.762 | 0.762 |

Notes: ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. The dependent variable is the share of anti-immigration votes with mean of 0.48 (sd = 0.15). See table 4.4 for information about the three groups of immigrants. Robust standard errors are given in parentheses and are clustered on municipality-level. All models use the baseline covariate and fixed effects specification.

Interestingly, it does not seem to make a difference whether the value difference is one- or two-dimensional. The point estimates are significant and with 1.12 and 1.13 virtually identical for both groups of culturally different immigrants. The first stage and reduced form coefficients are again reported in table A.3 in the appendix. Notice that the F-value of the excluded instrument for immigrants that differ in both value dimensions is 11 and thus just above the critical value. This result should therefore be interpreted with the necessary caution.

I also considered further decomposing the group of culturally similar immigrants. To be precise, I estimate separate effects for the largest ethnicities, i.e. Italians, Spaniards, and Germans. However, I fail to find a point estimate that is significantly different from zero for any of them.

4.5.3 The Role of the Voter Turnout

A fact that has been ignored so far is that voting is not compulsory. As a consequence, all the findings of the previous analysis are only based on individuals who participate in the voting process. Since there is evidence that this is a selected group of individuals, the question arises whether the effect of immigration is the same on the non-voting persons.²⁰ Unfortunately, there are neither observations whether a person participated nor what the non-participant would have voted had he or she been forced to.

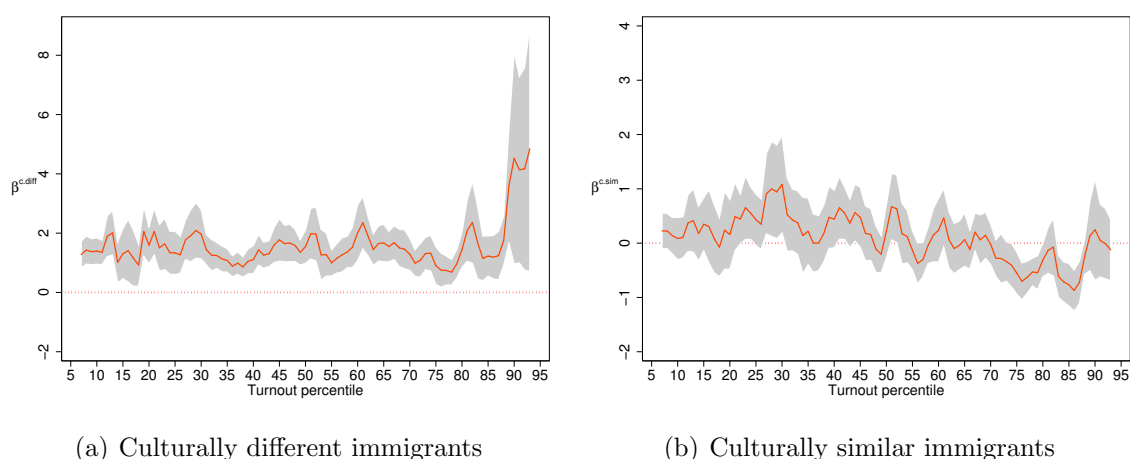
However, notice that the estimated effects would apply equally to all individuals if no unobserved characteristic were to exist that simultaneously influences anti-immigration attitudes and voting probability. Conversely, if there is a systematic relationship between participation decision and attitudes towards immigration that observed individual characteristics do not explain, one would expect to find a correlation between the size of the estimated effects and the voter turnout. I investigate this issue by re-running the baseline model for different turnout samples. In particular, I divide the overall sample into 100 subsamples, each containing observations (municipality vote cells) with a particular turnout percentile. I then estimate the effects for each subsample (turnout percentile) as depicted in figure 4.3.

The vertical axis indicates the impact of the culture-specific immigrant share (in percentage points), the horizontal axis depicts the corresponding turnout percentile. The overall variation in voter turnout is quite large. The median turnout corresponds to 47.8%, the 5th and the 95th percentile correspond to 26.5% and 72.3%, respectively.

²⁰For instance, according to the findings of Miguet (2008) who uses post-vote datasets (VOX survey) for two votes in Switzerland (votes 355 and 467), participation is significantly more likely for highly educated and married persons, and for individuals with views at the political extreme.

Consider the effect of culturally different immigrants first. Interestingly, the results do not provide any indication of a systematic relationship between the size of the effect and the voter turnout. For each turnout subsample, IV estimates are statistically significant and vary between 0.75 and 2.5 with exception of the upper decile. In those municipalities, the estimated effect is considerably larger (up to 4.75 percentage points), but so are the standard errors. Results for culturally similar immigrants again suggest no causal relationship between the share of culturally similar immigrants and voting results. These patterns are robust to a variety of checks with respect to the underlying subsamples.²¹ Also, there is no reason for weak instrument concerns.

Figure 4.3: The role of the turnout



Notes: These figures show the estimated effects (in pp) of the culture-specific immigrant share on anti-immigration votes depending on turnout (IV regressions). The grey shaded area depicts the 95% confidence interval. The effects are estimated for each turnout percentile using a moving average over two percentiles. All models use the baseline covariate and fixed effects specification.

In sum, the results suggest that no unobserved variable exists that jointly determines participation decision and attitudes towards immigration in a systematic manner. The effects thus appear to be independent of voter turnout once it is controlled for observed individual characteristics, which implies that even if voting were mandatory, the effects would be the same. It must be emphasized that this only applies to the estimated effects. In terms of the vote outcome, it may well be that forcing everybody to vote would change it. This is because there is reason to expect a systematic relationship between participation and voting decision. Probably the most well known characteristic that is associated with participation as well as voting decision is the level of education. On the one side, there is quite a consensus that the highly educated individuals are more likely to participate in the voting process (see

²¹For instance, I also divided the sample into 50 or 20 subsamples, respectively.

for Switzerland, for instance, Miguet (2008)). On the other side, higher education is found to be positively correlated with a more liberal mindset (see table A.2 in the Appendix). Forcing all individuals to vote would thus imply a relative decrease in liberal views and potentially result in a different outcome of the vote.

4.5.4 The Rise of Switzerland's Right-Wing Party

Similar to many other European countries, Switzerland has recently experienced a strong rise of the Swiss People's Party (SVP), a party that mainly attracts attention because of its fierce anti-immigration program.²² This last section aims at investigating whether the local presence of immigrants with different cultural values and beliefs is a similarly important determinant of the percentage of right-wing voters as it is of the percentage of anti-immigration votes. In other words, I am interested in the link between voting in favor of a more restrictive immigration policy and supporting the right-wing party.

To explore this question I draw on results of the parliamentary elections.²³ I apply the same estimation strategy as before, the only difference is the dependent variable, which now is the share of SVP votes instead of the share of anti-immigration votes. The sample is somewhat smaller than the sample that is used in the main analysis, the reason being the slightly different version of the municipality boundaries that is used in order to record election results. As a consequence, 24 of total 2544 municipalities cannot be merged to the census data.

Table 4.10 presents the results. The first two columns replicate the baseline results based on the slightly smaller sample. The last two columns show the impacts of the culture-specific immigrant shares on the proportion of SVP votes within a municipality. Besides the point estimates, the table also displays the change in the dependent variable in units of standard deviation as a result to a one standard deviation change in the main explanatory variable. These effects are expressed in square brackets.

²²From 1970 until 1991, the share of the votes for the SVP in parliamentary elections always lay at about 10 percent. In subsequent years, however, the party gained more and more support and reached its highest share of votes in 2007 with about 27 percent. This corresponds to an increase of almost three hundred percent over a period of just 16 years. The rapid rise of the SVP finally ended with the 2011 elections, where the party even had to record a slight decrease compared to the previous election period.

²³As the parliamentary elections take place every four years, the election years do not perfectly coincide with the census years. I therefore merge the 1971 election results to the 1970 census, the 1979 election results to the 1980 census, the 1991 election results to the 1990 census, the 1999 election results to the 2000 census and finally, the 2011 election results to the 2010 census.

Table 4.10: Additional results; Voting behavior and party affiliation

| | Dependent variable: share of | | | |
|--|------------------------------|----------|-----------|----------|
| | anti-immigration votes | | SVP votes | |
| | OLS | IV | OLS | IV |
| $I_{jT}^{\text{cult.diff.}}$ | 0.139** | 1.114*** | 0.165 | 0.910*** |
| | (0.066) | (0.217) | (0.105) | (0.174) |
| $\frac{\hat{\beta} \cdot sd(I_{jT}^{\text{c.diff}})}{sd(y)}$ | [0.053] | [0.423] | [0.058] | [0.323] |
| $I_{jT}^{\text{cult.sim.}}$ | -0.119*** | -0.201 | -0.046 | -0.164 |
| | (0.034) | (0.136) | (0.035) | (0.114) |
| $\frac{\hat{\beta} \cdot sd(I_{jT}^{\text{c.sim}})}{sd(y)}$ | [-0.059] | [-0.100] | [-0.023] | [-0.083] |
| F-stat. of excl. instr. ($I_{jT}^{\text{c.diff}}$) | | 47 | | 103 |
| F-stat. of excl. instr. ($I_{jT}^{\text{c.sim}}$) | | 139 | | 143 |
| Mean | 0.481 | 0.481 | 0.168 | 0.168 |
| Standard deviation | 0.152 | 0.152 | 0.153 | 0.153 |
| Number of observations | 57,960 | 57,960 | 12,601 | 12,601 |
| Adjusted R-Squared | 0.820 | 0.763 | 0.715 | 0.690 |
| p-value (F-statistic) | 0.000 | 0.000 | 0.000 | 0.000 |
| Area characteristics | Yes | Yes | Yes | Yes |
| Socio-demographic char. | Yes | Yes | Yes | Yes |
| NUTS2-specific vote FEs | Yes | Yes | Yes | Yes |

Notes: This table shows the effects of the culture-specific immigrant share on anti-immigration votes in the first two columns and on the share of SVP (Switzerland's far-right party) voters in the last two columns. ***, **, and * denote statistical significance on the 1%, 5%, and 10% level, respectively. All models use the baseline covariate and fixed effects specifications. Robust standard errors are given in parentheses and are clustered on municipality-level (2520 clusters).

The most striking aspect of the results is that the share of culturally different immigrants is found to affect the two dependent variables very similarly, not only qualitatively but also quantitatively. An increase in the local share of culturally different immigrants by one standard deviation leads to an increase in the proportion of anti-immigration votes by 0.4 standard deviations and to an increase in the share of right-wing voters by 0.3 standard deviations. There is obviously a tight link between voting and election behavior in connection with immigration, meaning that individuals who are inclined to vote for the right-wing party also vote in line with the party's program, or vice versa. There is no evidence with respect to culturally similar immigrants, however, that their presence causes people to vote for the far right.²⁴

²⁴These results are robust to all checks applied in table 4.8. The first stage and the reduced form coefficients are statistically significant, and the F-value of the excluded instrument lies well above the critical value.

These results compare well with the results reported by Halla *et al.* (2012) for Austria. For instance, they estimate that a one percentage point increase of the share of immigrants with a medium education level leads to an 1.1 percentage point increase in the share of FPÖ voters (the right-wing party in Austria).

4.6 Conclusions

This chapter combines census data with outcomes of national votes about immigration in Switzerland between 1970 and 2010 to analyze the causal relationship between the local immigrant share and attitudes to immigration. The focus lies in understanding how this relationship is affected by the level of dissimilarity in cultural values and beliefs between Swiss citizens and immigrants. The sorting of Swiss immigrants into different cultural groups is done based on findings from the World Value Surveys, covering all relevant source countries of Swiss immigrants. To take the endogenous settlement pattern of both, natives and foreigners into account, I instrument the municipality immigrant share with that of the local labor market. The identifying assumption is that the local labor market immigrant share is exogenously determined by family and friends or, put differently, that racially motivated mobility only exists within but not between local labor markets. I provide evidence that strongly supports the validity of this assumption.

The main finding of this chapter suggests that it is not the overall concentration of immigrants per se, but merely the presence of immigrants with a different cultural background that affects the voting behavior of Swiss citizens. More specifically, the local share of culturally different immigrants turns out to be a significant and sizable determinant of the percentage of anti-immigration votes. In contrast, the presence of culturally similar immigrants does not seem to affect natives' voting behavior at all. A one percentage point increase in the share of culturally different immigrants results in a 1.12 percentage point increase in anti-immigration votes in an average national vote about immigration. A wide range of robustness checks suggests that this result may be regarded as lower bound of the true effect. Moreover, checks show that it also holds for each single vote and that it is not sensitive to a finer cultural classification of immigrants. Additional analysis of participation rates suggests that the effects are not driven by unobserved characteristics of individuals who opt to participate in the voting process. The estimates in fact turn out to be independent of voter turnout if we control for observed individual characteristics. This implies that the main results of this study would still apply, even if voting had been mandatory. Finally, exploiting election results for the right-wing party SVP reveals a further

interesting finding. The share of culturally different immigrants appears to be an equally important determinant of the percentage of right-wing voters as it is for the percentage of anti-immigration votes, implying a tight link between Swiss citizens' voting behavior in plebiscites and referendums in comparison to elections.

One possible explanation for the sizeable positive effect of culturally different immigrants on anti-immigration attitudes may relate to cultural concerns (non-economic channel). More specifically, the native population may perceive culturally different immigrants as threatening to the national culture, language, religion, or way of life in general. However, since cultural values and beliefs are not the only difference between the two groups of immigrants, I cannot rule out the possibility that labor market and welfare concerns (economic channel) also explain part of the effect.

These findings may have implications for the continuing debates over immigration policy and help inform policy makers about possible impacts of the changing immigrant composition on outcomes of future votes. The suggestion that the ethnic composition of the locality in which individuals live might affect anti-immigrant attitudes may also have implications for continuing debates about the nature and location of new refugee accommodation centers.

4.A Appendix

Additional Tables and Figures

Table A.1: National Votes on Immigration Topics

| Nr. <i>v</i> | Year of vote, <i>t</i> | Dec- cade | Title of vote | Re- sult | Anti Imm. | Turn- out(%) | %Yes votes |
|-----------------|---------------------------|--------------|--|-------------|--------------|-----------------|---------------|
| 220 | 1970 | 1970 | PI “Against foreign infiltration” | rej. | yes | 74.7 | 46.0 |
| 242 | 1974 | 1970 | PI “Against foreign infiltration and overpopulation” | rej. | yes | 70.3 | 34.2 |
| 266 | 1977 | 1980 | PI “For the limitation of naturalizations” | rej. | yes | 45.2 | 33.8 |
| 310 | 1982 | 1980 | Modifications in the aliens act | rej. | no | 35.2 | 49.6 |
| 314 | 1983 | 1980 | FR on modifications of civil right regulations in the federal constitution | acc. | no | 35.8 | 60.8 |
| 315 | 1983 | 1980 | FR on easing certain naturalizations | rej. | no | 35.9 | 44.8 |
| 320 | 1984 | 1980 | PI “Against selling-off our homeland” | rej. | yes | 42.5 | 48.9 |
| 344 | 1987 | 1990 | Modifications in the asylum law | acc. | yes | 42.4 | 67.3 |
| 345 | 1987 | 1990 | FL on the stay and settlement of foreigners | acc. | yes | 42.2 | 65.7 |
| 355 | 1988 | 1990 | PI “For the limitation of immigration” | rej. | yes | 52.8 | 32.7 |
| 411 | 1994 | 1990 | FR on loosening naturalization regulations for young foreigners | rej. | no | 46.8 | 52.8 |
| 432 | 1996 | 2000 | PI “Against illegal immigration” | rej. | yes | 46.7% | 46.3% |
| 464 | 2000 | 2000 | FR on the bilateral agreements between Switzerland and the EU | acc. | no | 48.3 | 67.2 |
| 467 | 2000 | 2000 | PI “For an immigration regulation” | rej. | yes | 45.3 | 36.2 |
| 491 | 2002 | 2000 | PI “Against abuse of asylum right” | rej. | yes | 47.9 | 49.9 |
| 511 | 2004 | 2000 | FR on the acquisition of civil liberty for foreigners of the 3rd generation | rej. | no | 53.8 | 48.4 |
| 519 | 2005 | 2010 | FR on the expansion of the FMA to the new EU countries, and on the revision of the underpinning measures | acc. | no | 54.4 | 56.0 |
| 524 | 2006 | 2010 | FR on foreigners | acc. | yes | 49.0 | 68.0 |
| 525 | 2006 | 2010 | Change in the asylum law | acc. | yes | 49.1 | 67.8 |
| 532 | 2008 | 2010 | PI “For democratic naturalizations” | rej. | yes | 45.2 | 36.2 |
| 540 | 2009 | 2010 | FR on the permission to continue with the FMA’s and to additionally apply it to Bulgaria and Romania | acc. | no | 51.4 | 59.6 |
| 547 | 2009 | 2010 | PI “Against the building of minarets” | acc. | yes | 53.8 | 57.5 |
| 551 | 2010 | 2010 | PI “Deportation initiative” | acc. | yes | 52.9 | 52.9 |

Notes: This table lists all national votes on immigration topics between 1970 and 2010 whose percentage share of affirmative votes was between 30% and 70%. Abbreviations mean the following: FR: federal resolution, FL: federal law, PI: popular initiative; FMA: free mobility agreement.

Table A.2: Main Results, detailed

| | OLS | | IV (2SLS) MS-region-level instr. | |
|------------------------------------|----------------------|----------------------|-------------------------------------|----------------------|
| I_{jt} | 0.018 (0.027) | | 0.108 (0.139) | |
| $I_{jt}^{\text{cult.diff.}}$ | | 0.140** (0.066) | | 1.124*** (0.217) |
| $I_{jt}^{\text{cult.sim.}}$ | | -0.119*** (0.034) | | -0.201 (0.136) |
| <i>Local area characteristics:</i> | | | | |
| Proportion urban area | 0.058*** (0.013) | 0.058*** (0.012) | 0.059*** (0.013) | 0.059*** (0.014) |
| Border community | -0.012** (0.006) | -0.013** (0.006) | -0.011 (0.007) | -0.008 (0.008) |
| <i>Socio-demographic char's:</i> | | | | |
| Unemployed | -2.261*** (0.385) | -2.338*** (0.382) | -2.327*** (0.379) | -3.008*** (0.721) |
| Not employed | -0.457*** (0.078) | -0.465*** (0.077) | -0.443*** (0.088) | -0.429*** (0.114) |
| Not in working age | 0.395*** (0.086) | 0.353*** (0.085) | 0.462*** (0.135) | 0.539*** (0.147) |
| Self employed | 0.041 (0.035) | 0.048 (0.036) | 0.052 (0.033) | 0.130** (0.061) |
| Employed by family member | -0.134*** (0.046) | -0.176*** (0.045) | -0.114** (0.054) | -0.245*** (0.075) |
| Trainee | 0.855*** (0.077) | 0.795*** (0.077) | 0.875*** (0.081) | 0.660** (0.320) |
| Mgmt. position | -0.035 (0.063) | -0.039 (0.062) | -0.036 (0.063) | -0.065 (0.076) |
| Secondary education | -0.182*** (0.039) | -0.189*** (0.039) | -0.180*** (0.041) | -0.209*** (0.057) |
| Tertiary education | -0.784*** (0.043) | -0.787*** (0.044) | -0.752*** (0.065) | -0.636*** (0.096) |
| Single | -0.328*** (0.068) | -0.352*** (0.066) | -0.327*** (0.071) | -0.438*** (0.065) |
| Divorced, widowed | 0.119 (0.116) | 0.150 (0.111) | 0.010 (0.187) | -0.305 (0.325) |
| Age | 0.051*** (0.007) | 0.047*** (0.007) | 0.055*** (0.009) | 0.055*** (0.011) |
| Age ² | -0.001*** (0.000) | -0.001*** (0.000) | -0.001*** (0.000) | -0.001*** (0.000) |
| Women | -0.277** (0.113) | -0.220** (0.110) | -0.282** (0.119) | -0.016 (0.162) |
| Catholic | -0.044*** (0.007) | -0.043*** (0.007) | -0.047*** (0.008) | -0.051*** (0.010) |
| Jewish | 1.921** (0.827) | 1.844** (0.815) | 2.140** (1.050) | 2.697*** (1.041) |
| Islamic | 0.375 (0.496) | -0.265 (0.528) | 0.153 (0.730) | -4.090*** (1.018) |
| Other, no, unknown confession | -0.195*** (0.039) | -0.190*** (0.038) | -0.182*** (0.042) | -0.102 (0.068) |
| Born abroad | 0.229** (0.110) | 0.359*** (0.110) | 0.097 (0.242) | 0.185 (0.251) |

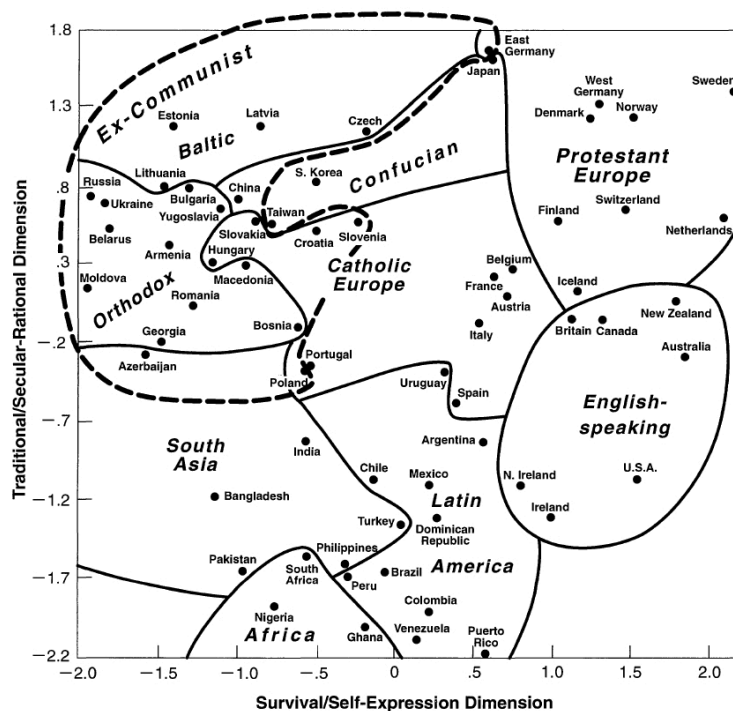
Notes: This table shows the detailed results of Table 4.7 for the baseline specification (M3). See Table 4.7 for further information. The baseline categories of the dummy-variables are defined as followed: employed, non mgmt.-position, mandatory education, married and evangelic.

Table A.3: First stage and reduced form coefficients

| Panel A: Main results | | | | | | |
|--|--------------------------|---------------------|---------------------------|---------------------------|----------------------|----------------------|
| | First stage coefficients | | | Reduced form coefficients | | |
| | M1 | M2 | M3 (baseline) | M1 | M2 | M3 (baseline) |
| I_{lt} | 1.001*** (0.060) | 0.442*** (0.036) | 0.389*** (0.040) | -0.437*** (0.054) | -0.069 (0.048) | 0.042 (0.053) |
| $I_{lt}^{cult.diff.}$ | 0.997*** (0.068) | 0.805*** (0.076) | 0.733*** (0.081) | 0.591** (0.272) | 1.026*** (0.094) | 0.896*** (0.096) |
| $I_{lt}^{cult.sim.}$ | 1.035*** (0.039) | 0.627*** (0.030) | 0.579*** (0.035) | -0.712*** (0.056) | -0.463*** (0.047) | -0.323*** (0.049) |
| Panel B: Robustness | | | | | | |
| | First stage coefficients | | | Reduced form coefficients | | |
| | M4 | M5 | M6 | M4 | M5 | M6 |
| $I_{lt}^{cult.diff.}$ | 0.668*** (0.058) | 0.514*** (0.035) | 0.639*** (0.153) | 0.982*** (0.070) | 1.300*** (0.053) | 0.345*** (0.093) |
| $I_{lt}^{cult.sim.}$ | 0.566*** (0.034) | 0.463*** (0.022) | 0.494*** (0.048) | -0.320*** (0.039) | -0.324*** (0.029) | -0.115* (0.062) |
| | M7 | M8 | M9 | M7 | M8 | M9 |
| | $x=j, 1970$ | $x=lt-1$ | $x=mt$ | $x=j, 1970$ | $x=lt-1$ | $x=mt$ |
| $I_x^{cult.diff.}$ | 0.822*** (0.178) | 0.694*** (0.110) | 0.765*** (0.113) | 0.717*** (0.187) | 0.614*** (0.141) | 1.732*** (0.124) |
| $I_x^{cult.sim.}$ | 0.435*** (0.023) | 0.454*** (0.039) | 0.448*** (0.039) | -0.026 (0.027) | -0.180*** (0.043) | -0.239*** (0.060) |
| Panel C: Sensitivity to cultural classification | | | | | | |
| | First stage coefficients | | Reduced form coefficients | | | |
| | M10 | | M10 | | | |
| $I_{lt}^{cult.sim.}$ | 0.592*** (0.034) | | -0.324*** (0.043) | | | |
| $I_{lt}^{c.diff.dim1}$ | 0.904*** (0.064) | | 0.875*** (0.109) | | | |
| $I_{lt}^{c.diff.dim2}$ | 0.955*** (0.336) | | 0.926*** (0.246) | | | |

Notes: This table shows the first stage and reduced form coefficients of the models estimated in table 4.7, table 4.8 and table 4.9. See the notes of these tables for further information.

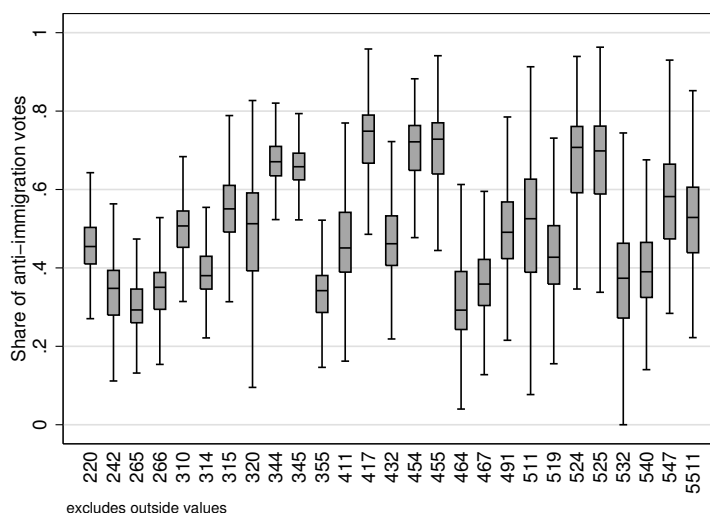
Figure A.1: Global cultural map



Source: Ronald Inglehart and Wayne E. Baker; Modernization, Cultural Change, and the Persistence of Traditional Values, *American Sociological Review*, **65**(1), page 29.

Notes: This figure shows the locations of 65 societies on two dimensions of cross-cultural variation: World Values Surveys, 1990-1991 and 1995-1998. The scales on each axis indicate the country's factor scores on the given dimension.

Figure A.2: Distribution of anti-immigration votes



Notes: This figure shows for each vote the distribution of the dependent variable, the share of anti-immigration votes at municipality-level, weighted by the corresponding population size.

Acknowledgments

I thank Armin Falk, Andrea Ichino, Andreas Kuhn, Kevin Lang, Rainer Winkelmann, Josef Zweimüller, as well as seminar participants in Zürich for helpful comments and suggestions. I also thank Kathrin Koller for great research assistance. The research was funded by the Austrian Science Fund (FWF): S 10304-G16.

Bibliography

- Alm, J. and Whittington, L. (1995). Does the income tax affect marital decisions? *National Tax Journal*, **48**, 565–572.
- Angrist, J. and Krueger, A. (1994). Why do World War II veterans earn more than nonveterans? *Journal of Labor Economics*, **12**(1), 74–97.
- Bauernberger, J. and Guger, A. (1998). Slight Decline of the Social Expenditure/GDP Ratio. Austria’s Social Expenditure in 1996. *Austrian Economic Quarterly*, **3**(3), 147–152.
- Bils, M. J. (1985). Real Wages over the Business Cycle: Evidence from Panel Data. *The Journal of Political Economy*, **93**(4), 666–689.
- Black, S., Devereux, P., and Salvanes, K. (2007). From the cradle to the labor market? The effect of birth weight on adult outcomes. *Quarterly Journal of Economics*, **122**(1), 409–439.
- Blanchflower, D. G. and Oswald, A. J. (1990). The Wage Curve. *The Scandinavian Journal of Economics*, **92**(2), 215–235.
- Blinder, A. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human Resources*, **8**(4), 436–455.
- Blundell, R., Reed, H., and Stoker, T. M. (2003). Interpreting aggregate wage growth: The role of labor market participation. *American Economic Review*, **93**(4), 1114–1131.
- Blundell, R., Francesconi, M., and Van der Klaauw, W. (2011). Anatomy of policy reform evaluation: Announcement and implementation effects. Technical report, Mimeo.
- Buckles, K. and Hungerman, D. (2008). Season of Birth and Later Outcomes: Old Questions, New Answers. NBER Working Papers No. 14573.

- Burgess, S., Propper, C., Rees, H., and Shearer, A. (2003). The class of 1981: the effects of early career unemployment on subsequent unemployment experiences. *Labour Economics*, **10**(3), 291–309.
- Card, D., Dustmann, C., and Preston, I. (2005). Understanding attitudes to immigration: The migration and minority module of the first european social survey. CReAM Discussion Paper Series No. 03.
- Chen, D. (2011). Can countries reverse fertility decline? evidence from france’s marriage and baby bonuses, 1929–1981. *International Tax and Public Finance*, **18**(3), 253–272.
- Cheng, D., Schwarz, E., Douglas, E., and Horon, I. (2009). Unintended pregnancy and associated maternal preconception, prenatal and postpartum behaviors. *Contraception*, **79**(3), 194–198.
- Clark, D. (2011). Do Recessions Keep Students in School? The Impact of Youth Unemployment on Enrolment in Post-compulsory Education in England. *Economica*, **78**(311), 523–545.
- Cnattingius, S. and Lambe, M. (2002). Trends in smoking and overweight during pregnancy: prevalence, risks of pregnancy complications, and adverse pregnancy outcomes. *Seminars in Perinatology*, **26**(4), 286–295.
- Devereux, P. J. (2000). Task assignment over the business cycle. *Journal of Labor Economics*, **18**(1), 98–124.
- Dickert-Conlin, S. and Chandra, A. (1999). Taxes and the Timing of Births. *Journal of Political Economy*, **107**(1), 161–177.
- Drago, R., Sawyer, K., Shreffler, K., Warren, D., and Wooden, M. (2011). Did Australia’s Baby Bonus Increase Fertility Intentions and Births? *Population Research and Policy Review*, **30**(3), 381–397.
- Dustmann, C. and Preston, I. (2001). Attitudes to ethnic minorities, ethnic context and location decisions. *The Economic Journal*, **111**(470), 353–373.
- Dustmann, C. and Preston, I. (2004). Is immigration good or bad for the economy? analysis of attitudinal responses. CReAM Discussion Paper Series No. 6.
- Dustmann, C. and Preston, I. (2007). Racial and economic factors in attitudes to immigration. *The BE Journal of Economic Analysis & Policy*, **7**(1), 62.
- Elsby, M. W., Hobijn, B., and Şahin, A. (2010). The labor market in the great recession. *Brookings Papers on Economic Activity*, **2010**(1), 1–48.
- Federal Ministry for Education, the Arts and Culture (2008). Development of Education in Austria, 2004–2007.
- Fernández, R. and Fogli, A. (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics*, pages 146–177.

- Fersterer, J., Pischke, J., and Winter-Ebmer, R. (2008). Returns to Apprenticeship Training in Austria: Evidence from Failed Firms. *Scandinavian Journal of Economics*, **110**(4), 733–753.
- Frühwirth-Schnatter, S., Pamminer, C., Weber, A., and Winter-Ebmer, R. (2011). Labor Market Entry and Earnings Dynamics: Bayesian Inference Using Mixtures-of-Experts Markov Chain Clustering. *Journal of Applied Econometrics*, **forthcoming**.
- Gans, J. and Leigh, A. (2009). Born on the first of July: An (un) natural experiment in birth timing. *Journal of Public Economics*, **93**(1-2), 246–263.
- Gardecki, R. and Neumark, D. (1998). Order from chaos? The effects of early labor market experiences on adult labor market outcomes. *Industrial and Labor Relations Review*, **51**(2), 299–322.
- Gathmann, C. and Schönberg, U. (2010). How General Is Human Capital? A Task-Based Approach. *Journal of Labor Economics*, **28**(1), 1–49.
- Genda, Y., Kondo, A., and Ohta, S. (2010). Long-Term Effects of a Recession at Labor Market Entry in Japan and the United States. *Journal of Human Resources*, **45**(1), 157–196.
- Gibbons, R. and Waldman, M. (2006). Enriching a theory of wage and promotion dynamics inside firms. *Journal of Labor Economics*, **24**(1), 59–107.
- Gnoth, C., Godehardt, D., Godehardt, E., Frank-Herrmann, P., and Freundl, G. (2003). Time to pregnancy: results of the german prospective study and impact on the management of infertility. *Human Reproduction*, **18**(9), 1959–1966.
- Hainmueller, J. and Hiscox, M. (2007). Educated preferences: Explaining attitudes toward immigration in europe. *International Organization*, **61**(02), 399–442.
- Hainmueller, J. and Hiscox, M. (2010). Attitudes toward highly skilled and low-skilled immigration: Evidence from a survey experiment. *American Political Science Review*, **104**(1), 61–84.
- Halla, M., Wagner, A., and Zweimüller, J. (2012). Does Immigration into Their Neighborhoods Incline Voters Toward the Extreme Right? The Case of the Freedom Party of Austria. IZA Discussion Paper No. 6575.
- Hugo, G. (2005). *Migrants in society: diversity and cohesion*. Global Commission on International Migration Geneva.
- Inglehart, R. and Baker, W. (2000). Modernization, cultural change, and the persistence of traditional values. *American sociological review*, **65**(1), 19–51.
- Inglehart, R. and Welzel, C. (2005). *Modernization, cultural change, and democracy: The human development sequence*. Cambridge University Press.
- Inglehart, R., Welzel, C., and Kligemann, H. (2003). The theory of human development: A cross-cultural analysis. *European Journal of Political Research*, **42**(3), 341–379.

- Jacobson, L. S., LaLonde, R. J., and Sullivan, D. G. (1993). Earnings losses of displaced workers. *The American Economic Review*, **83**(4), 685–709.
- Kahn, L. B. (2010). The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy. *Labour Economics*, **17**(2), 303–316.
- Kambourov, G. and Manovskii, I. (2009). Occupational Specificity of Human Capital. *International Economic Review*, **50**(1), 63–115.
- Kletzer, L. G. and Fairlie, R. W. (2003). The long-term costs of job displacement for young adult workers. *Industrial and Labor Relations Review*, **56**(4), 682–698.
- Kondo, A. (2008). Differential Effects of Graduating During a Recession Across Race and Gender. Mimeo, Columbia University.
- Kopczuk, W. and Slemrod, J. (2003). Dying to save taxes: Evidence from estate-tax returns on the death elasticity. *Review of Economics and Statistics*, **85**(2), 256–265.
- Kwon, I., Meyersson Milgrom, E. M., and Hwang, S. (2010). Cohort Effects in Promotions and Wages. Evidence from Sweden and the United States. *Journal of Human Resources*, **45**(3), 772–808.
- Lain, S., Ford, J., Raynes-Greenow, C., Hadfield, R., Simpson, J., Morris, J., and Roberts, C. (2009). The impact of the Baby Bonus payment in New South Wales: who is having “one for the country”? *Medical Journal of Australia*, **190**(5), 238–241.
- Liu, K., Salvanes, K. G., and Sørensen, E. Ø. (2012). Good skills in bad times: Cyclical skill mismatch and the long-term effects of graduating in a recession. IZA Discussion Paper No. 6820.
- Malani, A. and Reif, J. (2010). Accounting for anticipation effects: an application to medical malpractice tort reform. NBER Working Paper No. 16593.
- Mansour, H. (2009). The Career Effects of Graduating from College in a Bad Economy: The Role of Workers’ Ability. Mimeo, University of Colorado Denver.
- Martínez i Coma, F. and Duval Hernández, R. (2009). Hostility toward immigration in Spain. IZA Discussion Paper No. 4109.
- Mayda, A. (2006). Who is against immigration? a cross-country investigation of individual attitudes toward immigrants. *The Review of Economics and Statistics*, **88**(3), 510–530.
- Meuleman, B., Davidov, E., and Billiet, J. (2009). Changing attitudes toward immigration in Europe, 2002-2007: A dynamic group conflict theory approach. *Social Science Research*, **38**(2), 352–365.
- Miguet, F. (2008). Voting about immigration policy: What does the Swiss experience tell us? *European Journal of Political Economy*, **24**(3), 628–641.

- Moulton, B. R. (1986). Random Group Effects and the Precision of Regression Estimates. *Journal of Econometrics*, **32**(3), 385–397.
- Muehlemann, S., Wolter, S., and Wuest, A. (2009). Apprenticeship training and the business cycle. *Empirical Research in Vocational Education and Training*, **1**(2), 173–186.
- Mulder, E., Robles de Medina, P., Huizink, A., Van den Bergh, B., Buitelaar, J., and Visser, G. (2002). Prenatal maternal stress: effects on pregnancy and the (unborn) child. *Early Human Development*, **70**(1), 3–14.
- Müller, T. and Tai, S. (2010). Individual attitudes towards migration: a reexamination of the evidence. Mimeo.
- Murphy, K. M. and Welch, F. (1990). Empirical Age-Earnings Profiles. *Journal of Labor Economics*, **8**(2), 202–229.
- Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. *Journal of labor Economics*, **13**(4), 653–677.
- Neugart, M. and Ohlsson, H. (2011). Economic incentives and the timing of births: Evidence from the german parental benefit reform 2007. Mimeo, Technical University of Darmstadt.
- Neumark, D. (2002). Youth labor markets in the United States: Shopping around vs. staying put. *Review of Economics and Statistics*, **84**(3), 462–482.
- Nijkamp, P. and Poot, J. (2005). The last word on the wage curve? *Journal of Economic Surveys*, **19**(3), 421–450.
- Norbeck, J. and Tilden, V. (1983). Life stress, social support, and emotional disequilibrium in complications of pregnancy: a prospective, multivariate study. *Journal of Health and Social Behavior*, **24**(1), 30–46.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, **14**(3), 693–709.
- O’Connell, M. (2011). How do high-skilled natives view high-skilled immigrants? a test of trade theory predictions. *European Journal of Political Economy*, **27**(2), 230–240.
- OECD (2001). Trends in international migration 2001.
- Oreopoulos, P., von Wachter, T., and Heisz, A. (2012). The Short- and Long-Term Career Effects of Graduating in a Recession. *American Economic Journal: Applied Economics*, **4**(1), 1–29.
- Orr, S., Miller, C., *et al.* (1997). Unintended pregnancy and the psychosocial well-being of pregnant women. *Women’s Health Issues: Official Publication of the Jacobs Institute of Women’s Health*, **7**(1), 38–46.

- Orr, S., Miller, C., James, S., and Babones, S. (2008). Unintended pregnancy and preterm birth. *Paediatric and Perinatal Epidemiology*, **14**(4), 309–313.
- Oyer, P. (2006). Initial Labor Market Conditions and Long-Term Outcomes for Economists. *The Journal of Economic Perspectives*, **20**(3), 143–160.
- Oyer, P. (2008). The Making of an Investment Banker: Stock Market Shocks, Career Choice, and Lifetime Income. *The Journal of Finance*, **63**(6), 2601–2628.
- Parent, D. (2000). Industry-specific capital and the wage profile: Evidence from the national longitudinal survey of youth and the panel study of income dynamics. *Journal of Labor Economics*, **18**(2), 306–323.
- Pulley, L., Klerman, L., Tang, H., and Baker, B. (2002). The extent of pregnancy mistiming and its association with maternal characteristics and behaviors and pregnancy outcomes. *Perspectives on Sexual and Reproductive Health*, **34**(4), 206–211.
- Quillian, L. (1995). Prejudice as a response to perceived group threat: Population composition and anti-immigrant and racial prejudice in europe. *American Sociological Review*, **60**(4), 586–611.
- Raaum, O. and Røed, K. (2006). Do Business Cycle Conditions at the Time of Labor Market Entry Affect Future Employment Prospects? *The Review of Economics and Statistics*, **88**(2), 193–210.
- Rayl, J., Gibson, P., and Hickok, D. (1996). A population-based case-control study of risk factors for breech presentation. *American Journal of Obstetrics and Gynecology*, **174**(1), 28–32.
- Rodgers, J. and Udry, J. (1988). The season-of-birth paradox. *Biodemography and Social Biology*, **35**(3), 171–185.
- Scheve, K. and Slaughter, M. (2001). Labor market competition and individual preferences over immigration policy. *Review of Economics and Statistics*, **83**(1), 133–145.
- Schneider, S. (2008). Anti-immigrant attitudes in europe: Outgroup size and perceived ethnic threat. *European Sociological Review*, **24**(1), 53–67.
- Schwerdt, G., Ichino, A., Ruf, O., Winter-Ebmer, R., and Zweimüller, J. (2010). Does the color of the collar matter? Employment and earnings after plant closure. *Economics Letters*, **108**(2), 137–140.
- Semyonov, M., Raijman, R., and Gorodzeisky, A. (2006). The rise of anti-foreigner sentiment in european societies, 1988-2000. *American Sociological Review*, **71**(3), 426–449.
- Solon, G., Barsky, R., and Parker, J. A. (1994). Measuring the Cyclicity of Real Wages: How Important is Composition Bias. *The Quarterly Journal of Economics*, **109**(1), 1–25.

- Staiger, D. and Stock, J. (1997). Instrumental variables regression with weak instruments. *Econometrica*, **65**(3), 557–586.
- Stevens, K. (2007). Adverse Economic Conditions at Labour Market Entry: Permanent Scars or Rapid Catch-up? Mimeo, University College London.
- Sullivan, P. (2010). Empirical evidence on occupation and industry specific human capital. *Labour Economics*, **17**(3), 567–580.
- Tamm, M. (2009). The Impact of a Large Parental Leave Benefit Reform on the Timing of Birth around the Day of Implementation. Ruhr Economic Papers No. 98.
- Tolbert, C. and Hero, R. (1996). Race/ethnicity and direct democracy: An analysis of california’s illegal immigration initiative. *Journal of Politics*, **58**(3), 806–818.
- Topel, R. H. and Ward, M. P. (1992). Job Mobility and the Careers of Young Men. *Quarterly Journal of Economics*, **107**(2), 439–479.
- von Wachter, T. and Bender, B. (2008). Do Initial Conditions Persist between Firms? An Analysis of Firm-Entry Cohort Effects and Job Losers Using Matched Employer-Employee Data. In S. Bender, J. Lane, K. L. Shaw, F. Andersson, and T. von Wachter, editors, *The Analysis of Firms and Employees. Quantitative and Qualitative Approaches*, pages 135–162. National Bureau of Economic Research.
- von Wachter, T. and Bender, S. (2006). In the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers’ Careers. *The American Economic Review*, **96**(5), 1679–1705.
- Wozniak, A. (2010). Are College Graduates More Responsive to Distant Labor Market Opportunities? *Journal of Human Resources*, **45**(4), 944–970.
- Zweimüller, J., Winter-Ebmer, R., Lalive, R., Kuhn, A., Ruf, O., Wuellrich, J.-P., and Büchi, S. (2009). The Austrian Social Security Database (ASSD). NRN: The Austrian Center for Labor Economics and the Analysis of the Welfare State, Working Paper 0903.

Curriculum Vitae

Personal Information

Date of birth: 5. August 1980.

Nationality: Swiss.

Languages: German (native), English (fluent), French (good).

Education

- | | |
|---------------------|--|
| Sep 2008 – Jul 2013 | Doctoral Studies in Economics at the University of Zurich. |
| Jan 2009 – Jan 2010 | Graduate Studies in Microeconomics and Econometrics at the Study Center Gerzensee, Switzerland. |
| Nov 2001 – Sep 2008 | Studies in Economics with Major Finance at the University of Zurich. |
| Aug 1993 – Jan 2001 | Academic high school, Federal “Matura” Typus E (Economics), “Kantonsschule Züricher Oberland”, Wetzikon. |

Professional Experience

- | | |
|---------------------|---|
| May 2013 – today | Research Associate, Winterthur Institute of Health Economics, Zurich University of Applied Sciences. |
| Sep 2008 – Apr 2013 | Research associate, Department of Economics, University of Zurich. |
| Jan 2008 – Sep 2008 | Research Assistant, Institute for Empirical Research in Economics, University of Zurich. |
| Aug 2006 – Mar 2007 | Junior Analyst in the field of Corporate Finance Advisory at Quadrum Partners Ltd., Zurich (parallel studies, 50%). |
| Mar 2001 – Dec 2007 | UBS, Global Asset Management, Zurich and Chicago (full-time until Oct 2001, 40% afterwards). |